How to Practice Evidence-Based Psychiatry

Basic Principles AND Case Studies

Edited by
C. Barr Taylor, M.D.

How to Practice Evidence-Based Psychiatry

Basic Principles and Case Studies

This page intentionally left blank

How to Practice Evidence-Based Psychiatry

Basic Principles and Case Studies

Edited by

C. Barr Taylor, M.D.

American Psychiatric Publishing, Inc.

Washington, DC
London, England

Note: The authors have worked to ensure that all information in this book is accurate at the time of publication and consistent with general psychiatric and medical standards, and that information concerning drug dosages, schedules, and routes of administration is accurate at the time of publication and consistent with standards set by the U.S. Food and Drug Administration and the general medical community. As medical research and practice continue to advance, however, therapeutic standards may change. Moreover, specific situations may require a specific therapeutic response not included in this book. For these reasons and because human and mechanical errors sometimes occur, we recommend that readers follow the advice of physicians directly involved in their care or the care of a member of their family.

Books published by American Psychiatric Publishing, Inc., represent the views and opinions of the individual authors and do not necessarily represent the policies and opinions of APPI or the American Psychiatric Association.

If you would like to buy between 25 and 99 copies of this or any other APPI title, you are eligible for a 20% discount; please contact APPI Customer Service at appi@psych.org or 800-368-5777. If you wish to buy 100 or more copies of the same title, please e-mail us at bulksales@psych.org for a price quote.

Manufactured in the United States of America on acid-free paper
13 12 11 10 09 5 4 3 2 1
First Edition

Typeset in Adobe's Janson and Berthold's Poppl-Laudatio.

American Psychiatric Publishing, Inc.
1000 Wilson Boulevard
Arlington, VA 22209–3901
www.appi.org

Library of Congress Cataloging-in-Publication Data
How to practice evidence-based psychiatry : basic principles and case studies / edited by C. Barr Taylor. — 1st ed.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-58562-365-5 (alk. paper)
1. Evidence-based psychiatry. 2. Evidence-based psychiatry—Case studies. I. Taylor, C. Barr (Craig Barr)
[DNLM: 1. Psychiatry—methods. 2. Evidence-Based Medicine—methods. 3. Mental Disorders—therapy.
WM 140 H847 2010]
RC455.2.E94H69 2010

616.89—dc22

2009028112

British Library Cataloguing in Publication Data
A CIP record is available from the British Library.

Contents

PART II

Case Studies: Evidence-Based Medicine Applied to Major DSM-IV Disorders

PART III

Case Studies: Evidence-Based Psychiatric Practice in Different Settings

Appendices

Contributors

Bibi Das, M.D.
Staff Psychiatrist and Clinical Educator, Stanford University School of Medicine, Stanford, California

Ira D. Glick, M.D.
Professor of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Gregory E. Gray, M.D., Ph.D.
Clinical Professor of Psychiatry and the Behavioral Sciences, Keck School of Medicine, University of Southern California, Los Angeles, California; and private practice, Pasadena, California

Keith Humphreys, Ph.D.
Research Professor, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Terence A. Ketter, M.D.
Professor of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Lorrin M. Koran, M.D.
Clinical Professor of Psychiatry and Behavioral Sciences, Emeritus, Stanford University School of Medicine, Stanford, California

Ben Kurian, M.D., M.P.H.
Assistant Professor, Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, Texas

Anna Lembke, M.D.
Instructor, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Steven E. Lindley, M.D. Ph.D.
Assistant Professor, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California; Director of Outpatient Mental Health, VA Palo Alto Health Care System, Menlo Park, California

Gretchen Lindner, Ph.D.
Staff Psychologist, PTSD Clinical Team, VA Palo Alto Health Care System, Menlo Park, California

Joanna M. Marino, M.A.
Psychology Intern, University of Florida Department of Public Health; University of Florida, Gainesville, Florida

Anthony J. Mascola, M.D.
Assistant Clinical Professor, Department of Psychiatry and Behavioral Sciences, Stanford University Medical Center, Stanford, California

Matthew H. May, M.D.
Private Practice and Adjunct Clinical Faculty, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

James E. Mitchell, M.D.
Christoferson Professor and Chair, Department of Clinical Neuroscience; Chester Fritz Distinguished University Professor, University of North Dakota School of Medicine and Health Sciences; and President, Neuropsychiatric Research Institute, Fargo, North Dakota

Douglas S. Rait, Ph.D.
Chief, Couples and Family Therapy Clinic, Stanford Medical Center, and Clinical Professor of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California; Director, Family Therapy Program, VA Palo Alto Health Care System, Menlo Park, California

James Reich, M.D.
Adjunct Associate Professor of Psychiatry, Stanford University School of Medicine, Stanford, California; Clinical Professor of Psychiatry, University of California, San Francisco Medical School, San Francisco, California

Margaret F. Reynolds-May, B.A.
M.D. Candidate and Office of Student Affairs, Yale University School of Medicine, New Haven, Connecticut

Anthony J. Rothschild, M.D.
Irving S. and Betty Brudnick Endowed Chair and Professor of Psychiatry, Department of Psychiatry, and Director, Senior Scholars Program, University of Massachusetts Medical School, Worcester, Massachusetts

Richard E. Shanteau, M.D.
Staff Psychiatrist, Kaiser Foundation Hospitals, South San Francisco, California

Caroline Spiranovic, Ph.D.
Research and Development Associate, Sentiens Pty Ltd, West Perth, Australia

Violeta Tan, M.D.
Psychiatry Resident, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Dennis Tannenbaum, M.B.Ch.B., FRANZCP
Chief Executive Officer and Consultant Psychiatrist, Sentiens Pty Ltd, West Perth, Australia

C. Barr Taylor, M.D.
Professor of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Madhukar H. Trivedi, M.D.
Professor of Psychiatry, Betty Jo Hay Distinguished Chair in Mental Health, and Chief, Division of Mood Disorders, University of Texas Southwestern Medical School, Dallas, Texas

Po W. Wang, M.D.
Clinical Associate Professor, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Disclosure of Interests

The following contributors to this book have indicated a financial interest in or other affiliation with a commercial supporter, a manufacturer of a commercial product, a provider of a commercial service, a nongovernmental organization, and/or a government agency, as listed below:

Ira D. Glick, M.D.—*Grant/research support:* AstraZeneca, Bristol-Myers Squibb/Otsuka, GlaxoSmithKline, National Institute of Mental Health, Solvay, Shire, Pfizer, Eli Lilly and Company; *Consult/advisory board:* Impax, Bristol-Myers Squibb, Janssen, Eli Lilly and Company, Lunkbeck, Organon BioSciences, Pfizer, Shire, Solvay, Vanda; *Speakers' bureau:* AstraZeneca, Bristol-Myers Squibb/Otsuka, Janssen, Pfizer, Shire.

Terence A. Ketter, M.D.—*Grant/research support:* Abbott Laboratories, AstraZeneca, Bristol-Myers Squibb, Cephalon, Eli Lilly and Company, GlaxoSmithKline, Pfizer, Repligen Corporation, Wyeth. *Consultant:* Abbott Laboratories, AstraZeneca, Bristol-Myers Squibb, Dainippon Sumitomo Pharma, Eli Lilly and Company, GlaxoSmithKline, Janssen, Jazz Pharmaceuticals, Novartis, Organon International, Solvay, Valeant Pharmaceuticals, Vanda Pharmaceuticals, Wyeth, Xenoport. *Lecture honoraria:* Abbott Laboratories, AstraZeneca, Bristol-Myers Squibb, Eli Lilly and Company, GlaxoSmithKline, Noven Pharmaceuticals, Pfizer. *Employee* (Nzeera Ketter, M.D., spouse)*:* Johnson & Johnson.

Lorrin M. Koran, M.D.—*Grant/research support:* Eli Lilly and Company, Somaxon, Sepacor. *Consultant:* Jazz Pharmaceuticals. *Speakers' bureau:* Forest Laboratories.

Ben Kurian, M.D., M.P.H.—*Grant/research support:* Agency for Healthcare Research and Quality, National Institute of Mental Health, Targacept.

Anthony J. Rothschild, M.D.—*Grant/research support:* Cyberonics, National Institute of Mental Health, Novartis, Takeda, Wyeth. *Consultant:* Eli Lilly and Company, GlaxoSmithKline, Forest Laboratories, Pfizer. *Trademarks:* The Rothschild Scale for Antidepressant Tachyphylaxis (RSAT). *Other:* Eli Lilly and Company and Pfizer provided study medication free of charge for the National Institute of Mental Health Study of the Pharmacotherapy of Psychotic Depression (STOP-PD).

Alan F. Schatzberg, M.D.—*Consultant:* BrainCells, CeNeRx, Corcept (Co-Founder), Eli Lilly and Company, Forest Laboratories, Merck, Neuronetics, Novartis, Pathways Diagnostics, Pfizer, PharmaNeuroBoost, Quintlies, Synosis, Xytis, Wyeth. *Honorarium:* GlaxoSmithKline, Roche. *Patents:* Named inventor on pharmacogenetic use patents on prediction of antidepressant response.

Dennis Tannenbaum, M.B.Ch.B., F.R.A.N.Z.C.P.—CEO of Sentiens Pty, Ltd.

C. Barr Taylor, M.D.—*Grant/research support:* National Institute of Mental Health, National Institute on Aging, National Cancer Institute, Schweizerische Anorexia Nervosa Stiflung (Switzerland), Commonwealth Department of Health and Ageing (Australia), Department of Veterans Affairs.

Madhukar H. Trivedi, M.D.—*Grant/research support:* Agency for Healthcare Research and Quality, Corcept Therapeutics, Cyberonics, Merck, National Alliance for Research in Schizophrenia and Depression, National Institute of Mental Health, National Institute on Drug Abuse, Novartis, Pharmacia and Upjohn, Predix Pharmaceuticals, Solvay, Targacept. *Consultant/speaker:* Abbott Laboratories, Abdi Brahim, Akzo, AstraZeneca, Bristol-Myers Squibb, Cephalon, Fabre-Kramer Pharmaceuticals, Forest Laboratories, GlaxoSmithKline, Janssen, Johnson & Johnson, Eli Lilly and Company, Meade-Johnson International, Neuronetics, Parke-Davis Pharmaceuticals, Pfizer, Sepraco, VantagePoint, Wyeth-Ayerst Laboratories.

Po W. Wang, M.D.—*Grant/research support:* Abbott Laboratories, AstraZeneca, Bristol-Myers Squibb, Eisai, Elan Pharmaceuticals, Eli Lilly and Company, GlaxoSmithKline, Janssen, Novartis, Repligen Corporation, Shire, Solvay, Wyeth. *Consultant:* Abbott Laboratories, Corcept Therapeutics, Pfizer, Sanofi-Aventis. *Lecture honoraria:* AstraZeneca, Bristol-Myers Squibb, Eli Lilly and Company, GlaxoSmithKline, Pfizer, Sanofi-Aventis.

The following contributors stated that they had no competing interests during the year preceding manuscript submittal:

Bibi Das, M.D.
Gregory E. Gray, M.D., Ph.D.
Keith Humphreys, Ph.D.
Anna Lembke, M.D.
Steven E. Lindley, M.D., Ph.D.
Gretchen Lindner, Ph.D.
Joanna M. Marino, M.A.
Anthony J. Mascola, M.D.
Matthew H. May, M.D.
James E. Mitchell, M.D.
Douglas S. Rait, Ph.D.
James Reich, M.D.
Margaret F. Reynolds-May, B.A.
Richard E. Shanteau, M.D.
Caroline Spiranovic, Ph.D.
Violeta Tan, M.D.

Foreword

We are in a great practice era in psychiatry. Advances in both psychotherapy and psychopharmacology have led to several effective approaches to many mental health problems. And the knowledge base on which we base psychiatric practice only continues to grow. Thus, I am delighted with the publication of *How to Practice Evidence-Based Psychiatry: Basic Principles and Case Studies*. This valuable new book will help residents, practicing psychiatrists, and mental health workers find the most useful and relevant information to inform and improve their practices.

In this text, C. Barr Taylor, M.D., has approached expounding on evidence-based practice in two ways. First, he has done a masterful job in revising and updating a splendid book, the *Concise Guide to Evidence-Based Psychiatry*, by Gregory E. Gray, M.D., Ph.D., which provides details on how to obtain and interpret medical evidence. The revisions include new chapters on how to consider other sources of information, including guidelines and measurements, and how to provide effective long-term treatment to patients with complicated, comorbid problems. Second, Dr. Taylor has recruited experts from a variety of specialty areas and practice settings to describe how they incorporate the latest evidence and outcome studies into their practice. These cases are both interesting and inspiring. I trust you will find them relevant and helpful to your everyday practice.

Over the last few years, the American Psychiatric Association has developed tools to facilitate aspects of evidence-based psychiatric practice, including the publication of practice guidelines and performance measures. One example, the *Handbook of Psychiatric Measures*, provides information about and ready access to many measures relevant to diagnosis, treatment, and evidence-based psychiatric practice. American Psychiatric Publishing, Inc., also has published a variety of resource books that provide the latest evidence and information on available and developing treatments, including basic textbooks and practitioner guides to the core psychotherapy competencies. Still, as Dr. Taylor notes, the major challenge for the modern practitioner is to apply and incorporate into his or her practice the most effective psychosocial and biological interventions that are based on the latest evidence.

How to Practice Evidence-Based Psychiatry: Basic Principles and Case Studies is an invaluable tool in helping you become a better practitioner. My congratulations to Dr. Taylor and his colleagues.

Alan F. Schatzberg, M.D.

This page intentionally left blank

Preface

In recent years, there has been an increased emphasis on evidence-based psychiatric practice. Many practice settings encourage clinicians to follow evidence-based guidelines and algorithms, medical students and residents are expected to develop and retain skills in evidence-based practice, and clinicians are encouraged to follow procedures to carefully evaluate the quality of the evidence they use in their practice. Despite this emphasis, many psychiatrists and other mental health professionals remain unfamiliar with methods and philosophy of evidence-based medicine (EBM) and, even more important, have little experience or guidance as to how to use evidence and the products developed from evidence (guidelines, algorithms) in clinical practice. The purpose of this book is to address the latter issue. The application of evidence to real clinical cases is difficult. Much of the evidence comes from patients who are less impaired and complicated than those we see in practice, focuses on immediate rather than long-term issues, and fails to address the need for integrated treatment.

I start with the assumption that guidelines, algorithms, and other sources of evidence and the interpretation of evidence need to be adapted to the individual patient by the clinician. I also believe that the best way to practice evidence-based psychiatry is to set measurable treatment goals and then to monitor progress toward those goals.

This book was written with two major goals in mind. The first goal is to discuss the methods and philosophy of evidence-based psychiatry. To achieve this goal, I updated and expanded the *Concise Guide to Evidence-Based Psychiatry*, by Gregory E. Gray, M.D., Ph.D., a popular, well-written, and easily understandable book on this topic. I was fortunate to have Dr. Gray's help with this process. The second goal is to discuss how the psychiatrist and other mental health specialists can incorporate evidence-based psychiatry into their clinical practice. To do so requires clinicians to use various tools—some of which may not be strictly evidence based, such as guidelines, expert opinion, and their own clinical experience and expertise. I invited experienced clinicians, many of whom are experts in their own right, to discuss cases in which they followed aspects of evidence-based care. I also decided that it would be useful to discuss the use of evidence-based practice when American Psychiatric Association guidelines are available and in various settings, including private practice. As the reader will see, the case presenters approached this task in various ways.

This book was written with three audiences in mind. The first audience consists of psychiatrists and other mental health professionals who wish to learn about evidence-based psychiatry on their own and who wish to incorporate evidence-based psychiatry into their busy practice. The first section (the updated *Concise Guide*) can be used both as an introduction to the topic and as a ready reference for researching the literature and appraising evidence.

The second audience is psychiatry residents and other mental health trainees. Several textbooks on the topic of EBM are available, including the *Concise Guide to Evidence-Based Psychiatry*. The first section of this book, the updated *Concise Guide*, focuses on the needs of psychiatry residents and of other mental health trainees by emphasizing the information resources of most use in finding answers to clinical

questions in clinical practice. In addition, examples are drawn from the psychiatric literature rather than from general medicine or surgery. The second and third sections include cases that may be useful for residents. I have used these cases in my lectures on evidence-based psychiatry. I also invited a resident to discuss a patient she treated.

The third audience is residency program directors and faculty who are looking for a brief introduction to evidence-based psychiatry and examples of how evidence-based psychiatry can be practiced. The Accreditation Council for Graduate Medical Education requires that all residents develop EBM skills as part of the "practice-based learning and improvement" core competency, and the program directors and faculty need to be more aware of these practices. Chapter 15 includes updated suggestions from the *Concise Guide to Evidence-Based Psychiatry* for teaching EBM in psychiatry residency training programs.

It is my hope that the information and examples in this book can help achieve the goal of evidence-based psychiatry, which is to provide the best care to our patients.

C. Barr Taylor, M.D.

Cautionary Statement

The cases presented in this volume are meant to illustrate how experienced clinicians have approached various clinical issues. However, they are not meant to be recommendations as to how other clinicians should treat similar cases, and other clinicians might have equally useful approaches to their patients. The cases are not meant, either, to be recommendations of measures that should be used or dosages of medications that should be given unless their status as such is clearly stated by the authors.

Readers should note that Internet addresses change frequently. If a link provided is no longer current, a search from the referenced homepage by title or keyword may locate the cited document.

This page intentionally left blank

Acknowledgments

The development of *How to Practice Evidence-Based Psychiatry: Basic Principles and Case Studies* has relied on the help of many individuals. The first section of the book is primarily an update of the excellent *Concise Guide to Evidence-Based Psychiatry*, by Gregory E. Gray, M.D., Ph.D., with new chapters added to expand the focus to evidence-based psychiatry practice. I am deeply grateful for the chance to update this material. Dr. Gray was also very willing and available to review and revise updates.

I also would like to thank all the contributors to this book. All prepared interesting and important cases and responded to any suggestions for changes. I appreciate their willingness to be forthright in discussing how they approach cases with evidence-based medicine procedures.

Many individuals from American Psychiatric Publishing contributed in many important ways to this project. I particularly want to thank Bob Hales, Editor-in-Chief, and John McDuffie, Editorial Director, who provided much support and advice for this project, and Roxanne Rhodes, Senior Editor, and the Books Department staff for their careful work in editing and preparing the manuscript.

I want to thank the many individuals who read drafts of various chapters and provided important feedback and recommendations, including Vickie Chang, Darby Cunning, Debra Safer, Violeta Tan, and Mickey Trockel. This project began with discussions in the Behavioral Medicine Clinic in the Stanford Medical School Department of Psychiatry about how to practice evidence-based medicine. I want to thank Bruce Arnow, Chris Hayward, Elias Aboujaoude, Debra Safer, Anthony Mascola, and Kim Bullock for their suggestions about how to think about and present these difficult ideas. Smita Das helped me develop some of the original forms and procedures that eventually became central to the cases I present. Nicole DeBias was incredibly helpful in preparing, editing, and assembling the many pieces for submission to American Psychiatric Publishing.

This page intentionally left blank

I

Basic Principles of Evidence-Based Psychiatry

This page intentionally left blank

1

What Is Evidence-Based Psychiatric Practice?

C. Barr Taylor, M.D.
Gregory E. Gray, M.D., Ph.D.

Evidence-based psychiatric practice (EBPP) is a broad term referring to clinical practice that is informed by evidence about interventions and considers patient needs, values, and preferences and their integration in determining individual care (Kazdin 2008). EBPP uses evidence-based medicine (EBM) to assess the quality of evidence relevant to the risks and benefits of treatments (including lack of benefit). According to the Centre for Evidence-Based Medicine, "Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients" (Sackett et al. 1996).

Gray (2004), in *Concise Guide to Evidence-Based Psychiatry*, noted that EBM focuses on the use of clinical expertise to integrate research evidence, patient preferences, and clinical state in making decisions about patient care (Guyatt et al. 2000; Haynes et al. 2002a, 2002b; Straus and McAlister 2000). Usually, EBM addresses a particular issue (e.g., "What is the evidence that a particular treatment is effective for a particular problem?"). EBPP refers to the more complicated problem of integrating various sources of evidence, including experience, in treating patients in real-world settings and over long periods. The challenge of EBPP is to determine how "the evidence" fits a particular patient with multiple problems.

Over the last decade, the data from innumerable studies have been massaged by experts into guidelines, algorithms, consensus statements, reviews, textbooks, meta-analyses, and other documents. These are important sources for EBPP. In psychiatry, we also need to consider psychosocial, psychotherapeutic, and psychopharmacological sources of expertise and evidence. Thus, we think it is better to take a broad perspective and focus on both the *evidence* and the *expertise*. The goal of EBPP is to improve patient care, and we need to consider a set of practices relevant to achieving that aim.

Gray (2004) has noted that all practice guidelines involve some degree of judgment and bias in their development (Browman 2001; Drake et al. 2001; Greenhalgh 2001), the extent of which is often unstated. In addition, clinical practice guidelines are often introduced as part of a "top-down" approach to changing clinician behavior, which may lead to clinician resistance. Ideally, EBM and EBPP are "bottom-up" approaches, in which clinicians make decisions based on ability to search and appraise the medical literature and use this information in their practice situation. Yet in practice, few clinicians have the time to find and read systematic reviews and thus rely on guidelines and algorithms. What remains most important is that patients are provided the best treatment available, and the best way to determine that is to monitor progress against some standard. Taking these considerations together, EBPP includes

- Accurate evaluation and treatment planning
- Consideration of treatment algorithms, guidelines, and best practices when planning and providing treatment

- Use of measures to determine progress of an individual patient
- Review of a patient's progress against personal and published standards
- Consideration of "evidence and experts" in clinical decisions
- Expertise in providing a range of therapies or in making them available through other resources or providers
- Periodic review of general practice outcomes

The cases in Parts II and III of the book illustrate how various clinicians have considered treatment algorithms, guidelines, and best practices and measures when planning and delivering treatment. In Chapter 4, we discuss the many resources available to find evidence relevant to your clinical questions. Yet even with the availability of these many resources for facilitating EBPP, new knowledge will continue to be added, and the practitioner will be faced with the need to interpret the summary sources and to search for new information. In addition, the process an individual goes through in evaluating a source is similar to that used by many experts. For this reason, we have updated the *Concise Guide to Evidence-Based Psychiatry* (Gray 2004). Part I of the book can be used to learn the principles and practices of EBM and EBPP.

The Development of Evidence-Based Medicine and Psychiatry

EBM arose for a variety of reasons, but perhaps the main one was the need to find ways to use the vast knowledge accumulating from clinical trials in ways that might improve or provide more cost-effective care. A search (conducted 4/06/08) of the term "depression treatment" in PubMed identified more than 100,000 articles, with more than 10,000 published in the last few years. Many of the scientific practices of EBM are designed to find ways to systematically sort through such evidence to find information that might improve care.

EBM had its origin in the Department of Clinical Epidemiology and Biostatistics at McMaster University in Canada (Guyatt 2002). In 1981, members of that department began publishing a series of articles in the *Canadian Medical Association Journal* that were intended to teach clinicians how to critically appraise the medical literature. In 1990, they began to move beyond teaching critical appraisal skills and started developing a new philosophy of medical education, which they termed *evidence-based medicine*. In this new model, physicians would rely heavily on the medical research literature, rather than on textbooks or tradition, when approaching patient care problems. From 1993 to 2000, the McMaster group published a series of 25 articles in *JAMA* that defined many aspects of EBM. These articles have been revised and published in book form (Guyatt and Rennie 2002); they are also available, in their original form, at the Centre for Health Evidence (2007) Web site.

Although EBM was developed in Canada and described in *JAMA*, it has probably had its greatest effect in Britain. In part, this was because of the support of the National Health Service (NHS) (Baker and Kleijnen 2000; Ferguson and Russell 2000; Reynolds 2000; Trinder 2000). The NHS viewed EBM as a way both to improve the quality of care and to control costs by identifying and promoting therapies that worked and by eliminating therapies that were ineffective or harmful. The NHS funds university-based centers for EBM, surgery, child health, general practice, pathology, pharmacotherapy, nursing, dentistry, and mental health, as well as the NHS Centre for Reviews and Dissemination (1999) and the United Kingdom's Cochrane Centre (Baker and Kleijnen 2000; Sackett et al. 1996). In addition, the BMJ Publishing Group publishes several evidence-based journals and the semiannual publication *Clinical Evidence*, which is distributed to general practitioners by the NHS (Baker and Kleijnen 2000; Barton 2001).

In 1999, the Special Health Authority of the NHS in England and Wales established the National Institute for Clinical Excellence (NICE). In 2005, it joined with the Health Development Agency to become the new National Institute for Health and Clinical Excellence (still abbreviated NICE).

NICE publishes clinical appraisals of treatments with a consideration of cost-effectiveness. Since 2005, the NHS in England and Wales has been legally obliged to provide funding for medicines and treatments recommended by NICE's technology appraisal board. (Details as to how NICE develops guidelines can be found at www.nice.org.uk.)

NICE is not without its critics. As mentioned, it began with the goal of increasing access to effective treatments. It has been criticized, however, for the slow release of its appraisals; some of its assessments have seemed unfair, and the institute has been vili-

fied for recommendations to limit or deny coverage for some high-profile medicines for cancer and other life-threatening diseases. The most controversial decisions seem to be related to refusal to pay for expensive treatments assessed to have marginal benefit of extending life or improving quality of life. Some believe that one of the greatest threats of EBM is its use to determine what services will be reimbursed (Steinbrook 2008).

According to Gray (2002), the uptake of EBM has been slower in the United States than overseas. He noted that the slow incorporation of EBM into psychiatry practice may be related to mental health professionals' belief that their patients' individuality and the nonquantifiable aspects of psychotherapy preclude the application of EBM to psychotherapeutic interventions (Geddes 2000; Geddes et al. 1997; Goss and Rowland 2000; Kazdin 2008; Mace et al. 2001; Parry 2000). We hope the cases presented later in this book dispel this myth.

In recent years, however, EBPP has achieved wide influence in many systems. In the United States, the Agency for Healthcare Research and Quality (2008) has established the National Guideline Clearinghouse (NGC; www.guideline.gov). This important resource is a comprehensive database of evidence-based clinical practice guidelines and related documents. NGC was originally created by the Agency for Healthcare Research and Quality in partnership with the American Medical Association (AMA) and the American Association of Health Plans (now America's Health Insurance Plans). The NGC mission is to provide physicians, nurses, other health professionals, health care providers, health plans, integrated delivery systems, purchasers, and others an accessible mechanism for obtaining objective, detailed information on clinical practice guidelines and to further their dissemination, implementation, and use.

The Veterans Health Administration (VHA) of the U.S. Department of Veterans Affairs (VA) may be in the forefront of providing evidence-based mental health practice. The recent VHA Handbook lays out minimum requirements for VHA mental health services (U.S. Department of Veterans Affairs 2008). In the section on evidence-based psychotherapies, the VHA notes that all veterans with posttraumatic stress disorder (PTSD) must have access to cognitive processing therapy or prolonged exposure therapy and that medical centers and very large community-based outpatient clinics must provide adequate staff capacity to allow the delivery of evidence-based psychotherapy to their patients. In addition, all veterans with depression or anxiety disorders must have access to three evidence-based treatments, as appropriate: cognitive-behavioral therapy (CBT), acceptance and commitment therapy, or interpersonal therapy. In addition, all care sites need to provide evidence-based pharmacotherapy for mood disorders, anxiety disorders including PTSD, psychotic disorders, substance use disorders, dementia, and other cognitive disorders. The VHA noted, "Such care is to be consistent with current VA clinical practice guidelines and informed by current scientific literature." The VA has restricted internal Web sites providing access to VA clinical practice guidelines (vaww.oqp.med.va.gov/CPGintra/cpg/cpg.htm) and VHA clinical practice guidelines (vaww.national.cmop.va.gov/PBM/Clinical%20Guidance/Forms/AllItems.aspx).

EBPP also has become important in some large health maintenance organizations, such as the Kaiser Permanente Health Care System, and in many preferred provider organizations.

The Accreditation Council for Graduate Medical Education (2008), the body that sets standards for training residents in psychiatry and other specialties, requires training programs to show how they are teaching competence in the methods of EBM.

EBM has had a major influence on psychiatric practice in the United Kingdom because of the role of the NHS, as well as the efforts of the Centre for Evidence-Based Mental Health at the University of Oxford (Baker and Kleijnen 2000; Ferguson and Russell 2000; Geddes 2000). These efforts have been aided by the journal *Evidence-Based Mental Health*, which is published jointly by the BMJ Publishing Group, the Royal College of Psychiatrists, and the British Psychological Society (Geddes et al. 1997). In addition, skill in applying EBM is tested in the critical appraisal paper that is now included in part II of the membership examination of the Royal College of Psychiatrists (Dhar 2001; Geddes 2000).

In 2006, the United Kingdom established the Improving Access to Psychological Therapies program to provide evidence-based mental health treatment for patients with anxiety and depressive disorders (Department of Health and Care Services Improvement Partnership 2006). The ambitious goal of this program is to improve services to 900,000 or more

depressed people over the next few years, with half of those completing therapy recovering. The program uses a stepped-care model with the expectation that data will be obtained on all participants.

The American Psychological Association (2005) has been slow to promote evidence-based interventions, although they encourage evidence-based psychological practice. The Task Force on Evidence-Based Practice has a very nonspecific recommendation about evidence-based psychological practice:

> Clinical decisions should be made in collaboration with the patient, based on the best clinically relevant evidence, and with consideration for the probable costs, benefits, and available resources and options. It is the treating psychologist who makes the ultimate judgment regarding a particular intervention or treatment plan. The involvement of an active, informed patient is generally crucial to the success of psychological services. Treatment decisions should never be made by untrained persons unfamiliar with the specifics of the case. The treating psychologist determines the applicability of research conclusions to a particular patient. Individual patients may require decisions and interventions not directly addressed by the available research. The application of research evidence to a given patient always involves probabilistic inferences. Therefore, ongoing monitoring of patient progress and adjustment of treatment as needed are essential to EBPP [evidence-based psychological practice]. (p. 3)

Clinicians need to be aware of the many resources available to evaluate the effectiveness of psychological interventions. Several textbooks have recently been published on evidence-based psychological practice (e.g., Fisher and O'Donogue 2006; Freeman and Power 2007; Goodheart et al. 2006; Levy and Ablon 2009; Rubin 2007). The Oxford University Press has a division related to publishing detailed manuals for patients and therapists on how to apply evidence-based practices (www.us.oup.com/us/catalog/general/series/TreatmentsThatWork). One Web site (http://ucoll.fdu.edu/apa/lnksinter.html) provides links to more than 30 other Web sites providing data on evidence-based psychological interventions. The Substance Abuse and Mental Health Services Administration (SAMHSA) has created the National Registry of Evidence-Based Programs and Practices, a searchable database of interventions for the prevention and treatment of mental and substance use disorders (www.nationalregistry.samhsa.gov/).

Does Evidence-Based Medicine Improve Outcomes?

One should ask, however, whether the use of evidence improves outcomes. This is actually a difficult question to answer. It would not be ethical to conduct a controlled trial to determine whether the use of EBM guidelines, compared with nonuse, improved outcomes because the patients randomly assigned to the nonuse condition would be offered a treatment that other evidence would suggest to be inferior. One procedure is to estimate the benefit if guidelines were applied in a particular population or setting. For instance, some EBM guidelines have been identified to provide the "best care" for post–myocardial infarction (MI) patients, such as prescribing aspirin and beta-blockers. Following these guidelines has been shown to improve outcome compared with previous practice (Krumholz et al. 1998). For instance, Gemell et al. (2005) estimated that following the United Kingdom's National Service Framework guidelines for adopting the National Service Framework recommendations for pharmacological interventions would prevent an extra 1,027 deaths (95% confidence interval [CI] 418–1,994) in post–acute MI patients and an extra 37,899 (95% CI 25,690–52,503) deaths in heart failure patients in the first year after diagnosis. Lifestyle-based interventions would prevent an extra 848 (95% CI 71–1,614) deaths in post–acute MI patients and an extra 7,249 (95% CI 995–16,696) deaths in heart failure patients.

Collaborative care for depression treatment in populations is another example of an evidence-based quality improvement intervention. This model has been developed by Katon and others over many years (Katon and Seelig 2008). In a total of 37 randomized trials of collaborative care compared with usual primary care, collaborative care was associated with twofold increases in antidepressant adherence, improvements in depressive outcomes that last up to 5 years, increased patient satisfaction with depression care, and improved primary care satisfaction with treating depression (Katon and Seelig 2008). As another example, better adherence to guidelines developed to treat opioid-dependent patients in Veterans Affairs opioid substitution clinics resulted in

greater reductions in heroin and cocaine use and greater improvement in mental health (Trafton et al. 2007).

Does Evidence-Based Psychiatric Practice Improve Outcomes?

Examining the effects of EBM-specific practices on carefully defined populations and problems is different from studying EBPP (Kazdin 2008). Again, EBM focuses on discrete conditions and issues, and EBPP focuses on the broader care of patients. EBPP is not what researchers study (Kazdin 2008), and doing so is difficult. The parameters of EBPP have not been clearly defined. We gave examples in the previous section of how following guidelines can improve care, but following guidelines is only one aspect of EBPP as previously defined. The systematic application of the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study guidelines, as illustrated in Trivedi and Kurian, Chapter 21 of this volume, is likely to improve treatment of depression.

Psychiatric Clinical Practice Is Not Always Evidence Based

Over the past few years, evidence has accumulated to suggest that there is a significant gap between the knowledge obtained from clinical trials regarding effective treatments for mental disorders and the actual treatment received by patients in clinical practice (Drake et al. 2001; Lehman and Steinwachs 1998; U.S. Department of Health and Human Services 1999; Young et al. 2001). Similar discrepancies have been noted in other fields of medicine, in which practice often lags years behind research findings (Egger et al. 2001; Geyman 2000; Haines and Donald 1998; Haines and Jones 1994). Other examples are that only about half of patients with diabetes have hemoglobin A1c levels below the recommended level of 8.0%; only one-third of patients with hypertension receive adequate treatment to lower blood pressure below guideline-recommended levels. According to a recent review, among patients in the community and primary care settings, more than 50% of those with depression do not receive accurate diagnoses or any prescription of depression treatment. Of those who are prescribed treatment, more than 50% do not receive adequate dosages of antidepressants, and fewer than 10% receive evidence-based psychotherapy (Katon and Seelig 2008). Estimates suggest that about 40% of primary care patients drop out of depression treatment within 6 weeks, and fewer than half are taking antidepressants at 6 months (Simon 2002). Similar low levels of adherence to evidence-based guidelines have been found for other disorders such as substance abuse.

Wide variations are also seen in the way psychiatry and other medical specialties are practiced (Geddes and Harrison 1997; Geyman 2000). Clearly, some methods of treatment must be more effective than others and have more research to support their use, yet surveys conducted in academic medical centers have found that up to 40% of clinical decisions are unsupported by evidence from the research literature (Geddes et al. 1996; Greenhalgh 2001).

Why the Gap?

Two general types of information problems contribute to patients receiving less than optimal care (Gray 2002; Haynes et al. 1997). The first problem is one of "information overload," which creates difficulties for clinicians who want to determine which treatments are truly most effective. There are thousands of medical journals and millions of articles; therefore, no psychiatrist or other clinician should expect to keep up with all of the developments in his or her field. Furthermore, when one looks at the results of various studies, they often appear to be contradictory. In part, this is caused by false-positive and false-negative results, which often arise from small samples (Collins and MacMahon 2001; see also Gray and Taylor, Chapter 5 in this volume). One could consult review articles to summarize the literature, but most such reviews are "journalistic" or "narrative" reviews, not systematic reviews (see Gray, Chapter 6 in this volume). As a result, such articles are subject to the biases of the review's author(s), both in terms of studies cited and in the method of summarizing conflicting results (Cook et al. 1998; Egger et al. 2001; Greenhalgh 2001). Textbook chapters have the added problem of rapidly becoming out of date. All of this contributes to the lag before advances in treatment are recognized and find their way into practice.

The second type of information problem causes ineffective treatments to be adopted or maintained.

The problem is not a lack of information but rather the uncritical acceptance of available information. This may occur for a variety of reasons, such as overreliance on one's own clinical experiences or on expert opinion, the uncritical acceptance of results of single studies, and the excessive influence of pharmaceutical companies through advertising and sponsorship of speakers (Greenhalgh 2001; Sackett et al. 2000). At the same time, one's own clinical experiences and expert opinions are important sources of information on how to provide patient care.

Problems With Evidence-Based Medicine

There are also many good reasons to question evidence-based medicine. A common complaint we hear from clinicians is that randomized studies enroll patients very unlike "my patients." This statement has some truth to it. First, patients in clinical studies tend to be younger *or* older than many patients in clinical practice. In medicine (less so in psychiatry), women have been underrepresented in clinical trials. Minorities are often absent or present in such small numbers as to make subanalyses meaningless. Many clinical trials exclude 90% or more of potentially interested subjects. For instance, Yastrubetskaya et al. (1997) found that only 4% of 186 elderly patients with elevated Hamilton Rating Scale for Depression scores were eligible for a Phase III trial of a new antidepressant. Taylor et al. (2007) compared patients used in meta-analyses of cardiac rehabilitation programs with two large samples of patients actually in such programs. Compared with patients in actual practice, patients presented in the meta-analyses were younger (54 vs. 64 years), more likely to be males (90% vs. 74%), more likely to have had an MI (86% vs. 53%) and not a coronary artery bypass graft (6% vs. 24%), and more likely to be in the program longer (18 vs. 8 weeks).

Randomized studies often exclude comorbid problems, whereas most patients seen in clinical practice have one or more comorbidities. To determine whether the results from placebo-controlled studies conducted in patients with manic episode can be generalized to a routine population of hospitalized acute manic patients, Storosum et al. (2004) examined the baseline characteristics of 68 patients with 74 episodes of acute mania who had been referred for routine treatment. In this study only 16% of the manic episodes would have qualified for the hypothetical trial.

Exclusion criteria do not necessarily bias the study in terms of a favorable outcome, although they may limit the external validity of the study. For instance, in the Storosum et al. (2004) sample, the most common exclusion criterion was the use of contraceptives. Excluding such patients reduces the external validity of the trial. In contrast, some other exclusion criteria (e.g., comorbid alcohol and drug use disorders) may have resulted in an overestimation of the efficacy of antimanic medications. Another limitation is that many clinical trials are very short—lasting less than a few months—and do not deal with the many patients who do not improve. A notable exception, among others, is the STAR*D trial, which has attempted to apply evidence-based care to real clinical populations as mentioned earlier.

Faced with these issues, clinicians need to consider the evidence of a treatment's efficacy as a "source" of information. From our standpoint, the challenge of the clinician is to view all evidence with a keen eye to its limitations but a possibility for its usefulness. In the cases provided later in this book, we include examples of how clinicians have used—and struggled with—sources of evidence as they try to provide effective care. The evidence needs to fit the realities of the patient. Because of the clinical challenges of integrated evidence with real cases, we favor a focus on "how the patient is doing."

Evidence Is Often Not Available

EBM assumes that treatments proven to be efficacious should be used over those lacking in evidence. Unfortunately, depending on the criteria for effectiveness, very few treatments have been shown to be beneficial. For instance, the BMJ *Clinical Evidence* base has evaluated more than 2,500 treatments (www.clinicalevidence.bmj.com/ceweb/about/knowledge.jsp). According to their criteria, only 13% were rated as beneficial, 23% were rated as likely to be beneficial, 8% were rated as a trade-off between beneficial and harmful, 6% were rated as unlikely to be beneficial, and 4% were rated as likely to be ineffective or harmful. Of the treatments, 46%—the largest proportion—were of unknown effectiveness. Use of only clearly beneficial treatments would severely limit options in patient care. Furthermore, the evidence for such beneficial treatments is

likely to have been derived from studies that, as mentioned earlier, may have limited applicability to patients seen in clinical practice. A reasonable approach is first to consider treatments shown to be effective with similar patients and then to retain a healthy skepticism about other treatments.

Two particular situations in which lack of evidence is especially problematic are the head-to-head comparison of new pharmacological treatments and situations involving treatment-resistant patients. The first situation is problematic because new drugs are evaluated relative to placebos or older medications—not newer medications—in the Phase III trials conducted to obtain U.S. Food and Drug Administration approval. The second situation is problematic because few studies other than the STAR*D trial have systematically evaluated treatment approaches to treatment-resistant patients.

Problems in Practicing Evidence-Based Psychiatry

In addition to concerns about EBM and evidence-based treatments, several practical problems arise in practicing evidence-based psychiatry.

Limited Time to Look for Answers

The first problem is time. Finding new information relevant to a case can be time-consuming, and clinicians often have very tight schedules. It is important to consider that some general guidelines to help direct care can be easily obtained. The practitioner should periodically update guidelines and even customize them, as discussed in Taylor, Chapter 18 of this volume, if he or she believes that the evidence justifies doing so. Guidelines are also frequently updated and revised. Guidelines are available to cover most common conditions, and once they have been selected, using them requires less time.

A more complicated problem is how to build on these guidelines and algorithms, particularly to find evidence of therapies useful across procedures, and to evaluate *new* information.

Although many sources of evidence about treatments have been developed to aid the busy practitioner, such sources can be out of date. First, most EBM procedures rely on large data sets, and many meta-analyses may be focused on results from older studies and interventions. Thus, it is a challenge for the practitioner to evaluate a potentially useful new intervention. The procedures for doing so require following EBM procedures, as discussed in the chapters in Part I of this book, but can be time consuming. Many other sources of more easily obtained current information, such as industry-sponsored continuing medical education courses, information from pharmaceutical representatives, and industry-sponsored reviews and seminars, may be biased.

Lack of Incentives for Evidence-Based Psychiatric Practice

Some practice settings now incorporate best practices as one form of EBPP, and clinicians are held accountable to these standards. However, in most settings, clinicians are free to practice EBPP as they wish. EBPP, particularly because it involves some of the more structured psychotherapies, may feel more demanding than do less-structured, nondirected approaches. As noted earlier, indications are that reimbursement may become more related to EBPP. The main reason that one should practice EBPP is the belief that doing so will improve patient care.

No Resources for Assessment

EBPP requires assessment of progress. Although simple assessment procedures are available, more routine assessments and data management may seem too much of an encumbrance for many practitioners. We discuss some simple procedures that can help overcome some of these problems (Chapters 8, 14, 16, 17) and also provide models in which electronic systems can support EBPP (Chapters 21 and 33).

Summary

The goal of mental health professionals should be to provide the most effective treatments to their patients. Doing so requires that they can determine which treatment approaches are more likely to be effective for a given patient, and this is, in essence, what EBPP is all about. Over the last decade, the data from innumerable studies have been massaged by experts into guidelines, algorithms, consensus statements, reviews, textbooks, meta-analyses, and other documents. These resources and the general principles developed for EBM have the potential to improve psychiatric practice significantly.

References

Accreditation Council for Graduate Medical Education: General competencies. Available at: www.acgme.org/outcome/comp/compfull.asp. Accessed June 25, 2009

Agency for Healthcare Research and Quality: Evidence-based practice centers (AHRQ). Available at: http://www.ahrq. gov/clinic/epc. Accessed June 25, 2009

American Psychological Association: APA Policy Statement on Evidence-Based Practice. August 2005. Available at: http://www2.apa.org/practice/ebpstatement.pdf. Accessed December 2, 2008

Baker M, Kleijnen J: The drive towards evidence-based health care, in Evidence-Based Counseling and Psychological Therapies: Research and Applications. Edited by Rowland N, Goss S. Philadelphia, PA, Routledge, 2000, pp 13–29

Barton S: Using clinical evidence. BMJ 322:503–504, 2001

Browman GP: Development and aftercare of clinical guidelines: the balance between rigor and pragmatism. JAMA 286:1509–1511, 2001

Centre for Health Evidence: Users' Guides to Evidence-Based Practice. 2007. Available at: http://www.cche.net/usersguides/main.asp. Accessed June 20, 2008

Collins R, MacMahon S: Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. Lancet 357:373–380, 2001

Cook DJ, Mulrow CD, Haynes RB: Synthesis of best evidence for clinical decisions, in Systematic Reviews: Synthesis of Best Evidence for Health Care Decisions. Edited by Mulrow C, Cook D. Philadelphia, PA, American College of Physicians, 1998, pp 5–12

Department of Health and Care Services Improvement Partnership. Available at: http://www.iapt.nhs.uk. Accessed December 2, 2008

Dhar R: Evidence-based journal clubs and the critical review paper. Psychiatr Bull 25:67–68, 2001

Drake RE, Goldman HH, Leff HS, et al: Implementing evidence-based practices in routine mental health settings. Psychiatr Serv 52:179–182, 2001

Egger M, Smith GD, O'Rourke K: Rationale, potentials, and promise of systematic reviews, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 3–19

Ferguson B, Russell I: Towards evidence-based health care, in Evidence-Based Counseling and Psychological Therapies: Research and Applications. Edited by Rowland N, Goss S. Philadelphia, PA, Routledge, 2000, pp 30–43

Fisher JE, O'Donogue WE (eds): Practitioner's Guide of Evidence-Based Psychotherapy. New York, Springer, 2006

Freeman C, Power M (eds): Handbook of Evidence-Based Psychotherapies: A Guide for Research and Practice. New York, Wiley, 2007

Geddes J: Evidence-based practice in mental health, in Evidence-Based Practice: A Critical Appraisal. Edited by Trinder L, Reynolds S. Oxford, UK, Blackwell Scientific, 2000, pp 66–88

Geddes JR, Harrison PJ: Closing the gap between research and practice. Br J Psychiatry 171:220–225, 1997

Geddes JR, Game D, Jenkins NE, et al: What proportion of primary psychiatric interventions are based on evidence from randomised controlled trials? Qual Health Care 5:215–217, 1996

Geddes J, Reynolds S, Streiner D, et al: Evidence based practice in mental health. BMJ 315:1483–1484, 1997

Geyman JP: Evidence-based medicine in primary care: an overview, in Evidence-Based Clinical Practice: Concepts and Approaches. Edited by Geyman JP, Deyo RA, Ramsey SD. Boston, MA, Butterworth-Heinemann, 2000, pp 1–12

Goodheart CD, Kazdin AE, Sternberg RJ (eds): Evidence-Based Psychotherapy: Where Practice and Research Meet. Washington, DC, American Psychological Association, 2006

Goss S, Rowland N: Getting evidence into practice, in Evidence-Based Counseling and Psychological Therapies: Research and Applications. Edited by Rowland N, Goss S. Philadelphia, PA, Routledge, 2000, pp 191–205

Gray GE: Evidence-based medicine: an introduction for psychiatrists. J Psychiatr Pract 8:5–13, 2002

Gray GE: Concise Guide to Evidence-Based Psychiatry. Washington, DC, American Psychiatric Publishing, 2004

Greenhalgh T: How to Read a Paper: The Basics of Evidence Based Medicine, 2nd Edition. London, BMJ Books, 2001

Guyatt G: Preface, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp xiii–xvi

Guyatt G, Rennie D (eds): Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Chicago, IL, AMA Press, 2002

Guyatt GH, Haynes RB, Jaeschke RZ, et al: Users' guides to the medical literature, XXV: evidence-based medicine: principles for applying the users' guides to patient care. JAMA 284:1290–1296, 2000

Haines A, Donald A: Making better use of research findings. BMJ 317:72–75, 1998

Haines A, Jones R: Implementing findings of research. BMJ 308:1488–1492, 1994

Haynes RB, Sackett DL, Gray JA, et al: Transferring evidence from research into practice, II: getting the evidence straight. ACP J Club 126:A14–A16, 1997

Haynes RB, Devereaux PJ, Guyatt GH: Clinical expertise in the era of evidence-based medicine and patient choice. ACP J Club 136:A11–A14, 2002a

Haynes RB, Devereaux PJ, Guyatt GH: Physicians' and patients' choices in evidence based practice. BMJ 324:1350, 2002b

Katon WJ, Seelig M: Population-based care of depression: team care approaches to improving outcomes. J Occup Environ Med 50:459–467, 2008

Kazdin AE: Evidence-based treatment and practice: new opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. Am Psychol 63:146–159, 2008

Krumholz HM, Philbin DM Jr, Wang Y, et al: Trends in the quality of care for Medicare beneficiaries admitted to the hospital with unstable angina. J Am Coll Cardiol 31:957–963, 1998

Lehman AF, Steinwachs DM: Patterns of usual care for schizophrenia: initial results from the Schizophrenia Patient Outcomes Research Team (PORT) client survey. Schizophr Bull 24:11–20, 1998

Levy R, Ablon JS: Handbook of Evidenced-Based Psychodynamic Psychotherapy: Bridging the Gap Between Science and Practice. New York, Humana Press, 2009

Mace C, Moorey S, Roberts B (eds): Evidence in the Psychological Therapies: A Critical Guide for Practitioners. Philadelphia, PA, Brunner Routledge, 2001

National Health Service Centre for Reviews and Dissemination: Getting evidence into practice. Eff Health Care 5(1):1–16, 1999

Parry G: Evidence-based psychotherapy: an overview, in Evidence-Based Counseling and Psychological Therapies: Research and Applications. Edited by Rowland N, Goss S. Philadelphia, PA, Routledge, 2000, pp 57–75

Reynolds S: The anatomy of evidence-based practice: principles and methods, in Evidence-Based Practice: A Critical Appraisal. Edited by Trinder L, Reynolds S. Oxford, UK, Blackwell Scientific, 2000, pp 17–34

Rubin A: Practitioner's Guide to Using Research for Evidence-Based Practice. New York, Wiley, 2007

Sackett DL, Rosenberg WMC, Gray JAM, et al: Evidence-based medicine: what it is and what it isn't. BMJ 312:71–72, 1996

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Simon GE: Evidence review: efficacy and effectiveness of antidepressant treatment in primary care. Gen Hosp Psychiatry 24:213–222, 2002

Steinbrook R: Saying no isn't NICE—the travails of Britain's National Institute for Health and Clinical Excellence. N Engl J Med 359:1977–1981, 2008

Storosum JG, Fouwels A, Gispen-de Wied CC, et al: How real are patients in placebo-controlled studies of acute manic episode? Eur Neuropsychopharmacol 14:319–323, 2004

Straus SE, McAlister FA: Evidence-based medicine: a commentary on common criticisms. CMAJ 163:837–841, 2000

Taylor RS, Bethell HN, Brodie DA: Clinical trials versus the real world: the example of cardiac rehabilitation. British Journal of Cardiology 14:175–178, 2007

Trafton JA, Humphreys K, Harris AH, et al: Consistent adherence to guidelines improves opioid dependent patients' first year outcomes. J Behav Health Serv Res 4:260–271, 2007

Trinder L: Introduction: the context of evidence-based practice, in Evidence-Based Practice: A Critical Appraisal. Edited by Trinder L, Reynolds S. Oxford, UK, Blackwell Scientific, 2000, pp 1–16

U.S. Department of Health and Human Services: Mental health: a report of the Surgeon General. Washington, DC, U.S. Department of Health and Human Services, 1999

U.S. Department of Veterans Affairs: Uniform mental health services in VA medical centers and clinics. Available at: http://www.va.gov/vhapublications/ViewPublication.asp?pub_ID=1762 (p 31). Accessed June 25, 2009

Yastrubetskaya O, Chiu E, O'Connell S: Is good clinical research practice for clinical trials good clinical practice? Int J Geriatr Psychiatry 12:227–231, 1997

Young AS, Klap R, Sherbourne C, et al: The quality of care for depressive and anxiety disorders in the United States. Arch Gen Psychiatry 58:55–63, 2001

This page intentionally left blank

2

The 5-Step Evidence-Based Medicine Model

Gregory E. Gray, M.D., Ph.D.

An important aspect of evidence-based medicine (EBM) is being able to efficiently find accurate information and evidence relevant to clinical questions. The method and urgency of finding information vary from one setting to another. Evidence related to direct care of inpatients needs to be obtained quickly; clinicians usually have more time to obtain information relevant to outpatient practice. In this chapter, we focus on one approach to help clinicians answer clinical questions with primary evidence sources.

One EBM approach involves the application of a 5-step model (Table 2–1) to apply evidence from the medical literature to patient care problems (Gray 2002; Sackett et al. 2000). Details of this model are provided in Chapters 3–12. An example of the use of the technique in practice can be found in Chapter 18, "Major Depressive Disorder and Bulimia Nervosa," in this volume.

TABLE 2–1. The 5-step evidence-based medicine process

Step 1:	Formulate the question
Step 2:	Search for answers
Step 3:	Appraise the evidence
Step 4:	Apply the results
Step 5:	Assess the outcome

Step 1: Formulate the Question

The EBM process begins with a clinical question, which may involve issues related to the diagnosis, treatment, prognosis, or etiology of a patient's illness. As I describe in Chapter 3, the question is formatted to include a patient problem or diagnosis; the treatment, diagnostic test, risk factor, or prognostic factor of interest, as well as any comparison; and the outcome of interest.

Step 2: Search for Answers

After formulating the question, the next step is to try to find an answer in the literature. This step involves an assessment of the type of evidence that is most appropriate for answering the question as well as the actual search for the evidence. Details of both processes can be found in Gray and Taylor, Chapter 4 in this volume.

Step 3: Appraise the Evidence

After an article has been located, it is necessary to appraise its validity and importance before applying the results. The specific questions to ask about validity and importance depend on the type of study design and the nature of the question. In addition, clinicians must decide whether the results can be applied to their particular patients and in their setting. The appraisal process is discussed in detail in Chapters 5–7 and 9–13.

Step 4: Apply the Results to Your Patient

Assuming that the evidence that the clinician has found is valid, important, applicable to his or her patient, and feasible in the practice setting, the next step is to apply it to the care of the patient, which is where clinical expertise is most important.

Step 5: Assess the Outcome

Step 5 includes an evaluation of the clinician's performance in searching the literature and in finding an answer to the clinical question posed, as well as an assessment of the patient's response to treatment.

Some Shortcuts

Studies in academic settings have found that the full 5-step model can be incorporated into routine practice (Ball 1998; Del Mar and Glasziou 2001; Sackett et al. 2000). In nonacademic settings, however, practitioners frequently voice concerns that are related to lack of time and information resources, as well as to an inadequate knowledge of the EBM process (Ely et al. 2002; Haines and Donald 1998; Lipman 2000; McColl et al. 1998; Straus and McAlister 2000; Trinder 2000; Young and Ward 2001).

Several shortcuts can be taken to streamline the process and to make it more practical in everyday clinical practice. First, it is important to recognize that a clinician does not have to go through the 5-step process for every patient encounter. After a question has been researched for one patient with a particular diagnosis, the answer can be applied to similar patients with the same diagnosis. In addition, because most patients of most clinicians fall into relatively few diagnostic categories, it soon becomes the exceptional patient who triggers the application of the full 5-step process. However, we believe it is important for the practitioner to set aside some time on a routine basis to search the literature and review evidence.

One way to keep up with the literature is to use preappraised information resources. Reviews, synopses, and summaries are available that address many issues raised by clinicians. The *BMJ* group has developed a whole set of tools to foster EBM (http://ebmh.bmj.com.). Some of these are available for free on the Internet, and others can be purchased. Most practitioners now have access to important databases through the Internet. In Chapter 4, we discuss how to assess these resources.

The use of structured approaches to asking and answering questions can be an excellent approach to teaching EBM.

For instance, the Chairman's Rounds at the Department of Psychiatry at Duke University has residents use the QUEST model (QUestion, Evaluate, and SynThesize) to address a clinical issue. For instance, one of the grand rounds addressed the following issues: How do long-term medication and behavioral treatments compare with each other? Are there additional benefits when they are used together? What is the effectiveness of systematic, carefully delivered treatments compared with routine community care? The resident then provided a critical review of a randomized trial that addressed that issue, which included a review of the study design, with an assessment of validity, a discussion of strengths and weaknesses, and the bottom line. The grand rounds topics, indexed by diagnosis, can be found at http://psychiatry.mc.duke.edu/Residents/Quest.html. Despite these shortcuts, there still will be questions that cannot be readily answered by the results of a previous search (or by the use of preappraised information resources). It is important, therefore, that psychiatrists and other clinicians be able to carry out the full 5-step EBM process when necessary (Evans 2001; Gray 2002; Guyatt et al. 2000; R.B. Haynes 2001; Straus et al. 2002).

References

Ball C: Evidence-based medicine on the wards: report from an evidence-based minion. Evid Based Med 3:101–103, 1998

Del Mar CB, Glasziou PP: Ways of using evidence-based medicine in general practice. Med J Aust 174:347–350, 2001

Ely JW, Osheroff JA, Ebell MH, et al: Obstacles to answering doctors' questions about patient care with evidence: qualitative study. BMJ 324:710–713, 2002

Evans M: Creating knowledge management skills in primary care residents: a description of a new pathway to evidence-based practice in the community. Evid Based Med 6:133–134, 2001

Gray GE: Evidence-based medicine: an introduction for psychiatrists. J Psychiatr Pract 8:5–13, 2002

Guyatt GH, Meade MO, Jaeschke RZ, et al: Practitioners of evidence based care. BMJ 320:954–955, 2000

Haines A, Donald A: Getting research findings into practice: making better use of research findings. BMJ 317:72–75, 1998

Haynes RB: Of studies, summaries, synopses, and systems: the "4S" evolution of services for finding current best evidence. Evid Based Ment Health 4:37–39, 2001

Lipman T: Evidence-based practice in general practice and primary care, in Evidence-Based Practice: A Critical Appraisal. Edited by Trinder L, Reynolds S. Oxford, UK, Blackwell Scientific, 2000, pp 35–65

McColl A, Smith H, White P, et al: General practitioners' perception of the route to evidence based medicine: a questionnaire study. BMJ 316:361–365, 1998

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Straus SE, McAlister FA: Evidence-based medicine: a commentary on common criticisms. CMAJ 163:837–841, 2000

Straus S, McAlister F, Cook D, et al: Expanded philosophy of evidence-based medicine: criticisms of evidence-based medicine, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 211–222

Trinder L: A critical appraisal of evidence-based practice, in Evidence-Based Practice: A Critical Appraisal. Edited by Trinder L, Reynolds S. Oxford, UK, Blackwell Scientific, 2000, pp 212–241

Young JM, Ward JE: Evidence-based medicine in general practice: beliefs and barriers among Australian GPs. J Eval Clin Pract 7:201–210, 2001

This page intentionally left blank

3

Asking Answerable Questions

Gregory E. Gray, M.D., Ph.D.

Every clinical encounter generates questions. Some of these require information that can be obtained only from the patient or from a collateral source, such as a friend or family member. Answers to such questions are unique to the particular patient and concern that patient's illness or situation. These types of questions form the basis of a psychiatric interview and are not the subject of this chapter.

Background Questions Versus Foreground Questions

In this chapter, I focus on questions about a patient's illness that are more general and hence answerable in the psychiatric literature. Such questions are often divided into two categories: background and foreground questions (McKibbon et al. 2002; Sackett et al. 2000).

Background questions concern relatively well-established facts. These are the sorts of questions that are best answered by reference books or textbooks. Examples of background questions are:

- What are the DSM-IV-TR diagnostic criteria for panic disorder?
- What are the dosage forms for olanzapine?
- What is cognitive-behavioral therapy?

Such questions typically have two parts: the journalistic "who, what, when, where, why, and how" and the name of a disorder or therapy. Background questions are the types of questions that medical students and beginning residents most frequently ask.

Foreground questions, in contrast, concern the current best information on diagnosis, treatment, or prognosis of a disorder. Such questions are best answered by the research literature, not by textbooks, because they involve information that is still in a state of flux as new knowledge is accumulated. Foreground questions are the most frequent type of questions generated by senior residents or by practicing clinicians.

The 4-Part PICO Question

Foreground questions are best framed as a 4-part question (Table 3–1). The question should include the patient(s) or problem of interest; the intervention of interest, including any comparison group; and the outcomes of interest (Badenoch and Heneghan 2002; Dawes 1999; Geddes 1999; McKibbon et al. 2002; Sackett et al. 2000). Such questions are sometimes referred to as PICO questions (i.e., questions that use the mnemonic aid "**p**atient/problem, **i**ntervention, **c**omparison, and **o**utcome") (Gray 2002). A question formulated in this way provides the parameters needed to conduct an efficient literature search because it makes the type of information required very clear.

Patients or Problem

The first part of the question involves the patients or problem of interest. The degree of specificity has an influence on the ability to find an answer, as well as on the applicability of that answer to the particular patient of interest. Because the 4-part question is

TABLE 3–1. The 4-part PICO question

P:	Patient or problem of interest
I:	Intervention of interest
	Treatment
	Diagnostic test
	Risk factor
	Prognostic factor
C:	Comparison
	Implicit or explicit
O:	Outcome of interest
	Positive or negative

used as the starting point for the literature search, the patient or problem should be defined with a degree of specificity consistent with the way that study populations are defined. For example, if the clinician is interested in the treatment of depression in preadolescents, specifying that population as "preadolescents with major depression" is preferable to "patients with major depression" because there are reasons to believe that the response to treatment might be different. However, if the clinician becomes overly specific and specifies "9-year-old Latino girls with a first episode of major depression and with a paternal but not maternal history of recurrent major depression," he or she may not be able to find evidence that is specific to such a narrowly defined population.

Intervention and Comparison

Intervention can be a treatment or a diagnostic test. Loosely defined, it can also refer to a risk factor or prognostic factor. In most cases, the intervention of interest will be compared with another intervention; in some cases, the comparison is explicit.

For questions related to treatment, the intervention can be either a specific medication or a psychosocial intervention. The comparison can be another active treatment, a placebo, or "usual care."

Diagnostic questions are usually questions related to the performance of a diagnostic test, screening instrument, or rating scale. The performance is usually compared with either a diagnostic "gold standard" or outcome on a commonly used instrument. Details of such comparisons are given in Chapter 9, "Diagnostic Tests," in this volume.

For questions related to etiology, the "intervention" is actually a risk factor. Here the comparison is often implicit (i.e., the absence of that risk factor).

Questions related to prognosis may be either 3- or 4-part questions. A 3-part question might ask the prognosis of patients with first-episode schizophrenia in general, whereas a 4-part question might ask whether a particular patient characteristic (prognostic factor) alters the prognosis. Here, as in questions related to etiology, the comparison can be implicit (i.e., the absence of the particular prognostic factor).

Outcome(s)

The fourth part of the question relates to the outcome or outcomes of interest. Such outcomes can be either positive (clinical improvement, remission, or survival) or negative (relapse, self-injury, or death). For questions involving diagnosis, the outcome is a measure of agreement between the two diagnostic methods.

Examples of 4-Part PICO Questions

Treatment

Example of a 4-part question related to treatment:

> In adult patients with schizophrenia, does the addition of cognitive-behavioral therapy to usual care, compared with usual care alone, prevent relapse?

In this example, the patients are "adult patients with schizophrenia," the intervention is the "addition of cognitive-behavioral therapy to usual care," the comparison is "usual care alone," and the outcome is "relapse."

Diagnosis

Example of a 4-part question related to diagnosis:

> In a primary care clinic population, is the self-administration of a brief screening instrument as effective in identifying patients with major depression as is a structured brief clinical interview by a psychiatrist?

In this example, the patients are "a primary care clinic population," the intervention (diagnostic test) is a self-administered "brief screening instrument," the comparison (i.e., the "gold standard") is "a structured brief clinical interview by a psychiatrist," and the outcome is "a diagnosis of major depression."

Etiology/Harm

Example of a 4-part question related to etiology:

> Among rescue workers at the site of the World Trade Center disaster, does the amount of time working at the disaster site influence the risk of developing posttraumatic stress disorder?

In this example, the population of interest is "rescue workers at the site of the World Trade Center disaster," the intervention (risk factor) is "the amount of time working at the disaster site," the comparison is implicit (i.e., less time vs. more time), and the outcome is the onset of "posttraumatic stress disorder."

A different type of question related to harm, this time with an explicit comparison group:

> In elderly patients (age 65 years and older) with schizophrenia receiving maintenance therapy with an antipsychotic medication, are patients receiving olanzapine at less risk for developing tardive dyskinesia than are patients receiving risperidone?

In this example, the patients are "elderly patients (age 65 years and older) with schizophrenia," the treatment of interest is "olanzapine," the comparison treatment is "risperidone," and the outcome of interest is the "development of tardive dyskinesia."

Prognosis

Example of a 4-part question related to prognosis:

> In patients ages 15–45 years, do female patients have better social and vocational functioning than do male patients after experiencing their first episode of schizophrenia?

In this example, the patients are persons "ages 15–45 years…after experiencing their first episode of schizophrenia," the intervention (prognostic factor) is being "female," the comparison group is "male patients," and the outcome is "social and vocational functioning." Had the question not explicitly asked about differences between male and female patients, it would have been an equally acceptable 3-part prognostic question.

References

Badenoch D, Heneghan C: Evidence-Based Medicine Toolkit. London, BMJ Books, 2002

Dawes M: Formulating a question, in Evidence-Based Practice: A Primer for Health Professionals. Edited by Dawes M, Davies P, Gray A, et al. New York, Churchill Livingstone, 1999, pp 9–13

Geddes J: Asking structured and focused clinical questions: essential first step of evidence-based practice. Evid Based Ment Health 2:35–36, 1999

Gray GE: Evidence-based medicine: an introduction for psychiatrists. J Psychiatr Pract 8:5–13, 2002

McKibbon A, Hunt D, Richardson WS, et al: Finding the evidence, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 13–47

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

This page intentionally left blank

4

Searching for Answers

Gregory E. Gray, M.D., Ph.D.
C. Barr Taylor, M.D.

After formulating a question, the next step in the evidence-based medicine (EBM) process is to search for the best evidence to answer it. However, before beginning a search for an answer to a clinical question, it is important to understand the nature of the question and the type of evidence that would best address it.

What Type of Evidence Is Best?

Some experts propose that certain types of research results should be given more weight than other types (Badenoch and Heneghan 2002; Phillips et al. 2001). In this model, types of evidence are put on a hierarchy with systematic reviews of randomized controlled trials (RCTs) being the best source of evidence and expert opinion the worst (see Table 4–1). The specific hierarchy depends on the type of clinical question being asked. Table 4–1 presents the evidence hierarchy for studies of therapy or harm. Types of evidence higher in the hierarchy are more apt to give a valid and unbiased estimate of the effect of an intervention than are those lower in the hierarchy. When attempting to answer a clinical question, the clinician should always rely on the evidence found that is highest in the hierarchy or in other ways determined to be the best evidence available.

These hierarchies of evidence have been validated by comparing results obtained from studies that addressed the same question but used different designs. In studies of a variety of therapies, for example, it has repeatedly been shown that observational studies may give misleading results compared with RCTs (Lacchetti and Guyatt 2002). Whether the observational studies over- or underestimate the effectiveness of a particular treatment or preventive practice depends on the specific intervention being considered (Reeves et al. 2001).

As an example of the use of an evidence hierarchy for questions related to treatment (Table 4–1), results from a systematic review of RCTs should be given more weight than the results from a single RCT, and results from an RCT should be given more weight than results from uncontrolled or nonexperimental studies. Thus, if a literature search produces a systematic review of RCTs, several individual RCTs, case reports, and an editorial, the clinician should rely on the systematic review because it is likely to be highest in the evidence hierarchy. However, one well-designed and conducted clinical trial with an adequate number of subjects followed up for a long time might very well provide much better evidence than that provided by analyzing results from many poorly conducted, short-term trials.

Expert Opinion

In the original publication (Gray 2004), expert opinion was listed, in Table 4–1, below case series as the worst source of evidence on the basis that some studies have suggested that expert opinion does not necessarily reflect the best evidence found in the literature (Antman et al. 1992) and *may* be influenced by personal gain or benefit from the recommendations. On the other hand, experts are often

TABLE 4–1. Hierarchy of evidence for studies of therapy or harm

Quality	Type of evidence
1a (best)	Systematic review of RCTs
1b	Individual RCT with narrow confidence interval
1c	"All-or-none" case series
2a	Systematic review of cohort studies
2b	Individual cohort study
	RCT with <80% follow-up
2c	Outcomes research
	Ecological study
3a	Systematic review of case-control studies
3b	Individual case-control study
4 (worst)	Case series

Note. RCT = randomized controlled trial.
Source. Adapted from Gray 2004; Phillips et al. 2001; Sackett et al. 2000.

used to write systematic reviews, perform meta-analyses, construct guidelines, and recommend best practices. Thus, expert opinion influences all sources of information and is a valuable and necessary source of information on how to choose treatments. Expert opinion should not, however, be considered fact unless the expert has provided strong evidence for that opinion. Furthermore, the expertise and bias of the experts should be considered in judging their recommendations in all sources of evidence-based psychiatry practice (EBPP).

Experts are considered to be individuals with considerable knowledge about a subject. Experts are called on to write guidelines, reviews, and synopses and to provide opinions about treatments. In judging expert opinion, one must consider both the person's level of expertise and how biased his or her judgment is likely to be about the issue at hand. A biased opinion is not necessarily bad. An expert's strong conviction about a treatment he or she has studied or developed may be appropriate. However, the expert must reassure the reader that his or her interpretation of the evidence base is influenced by this bias.

In medical publishing, journals now require authors to declare any conflicts of interest they might have. A conflict of interest is often defined as an actual or a perceived interest by an expert in an action that results in, or has the appearance of resulting in, personal, organizational, or professional gain (www.icmje.org/#conflicts). The assumption is that the declaration of conflict of interest allows the reader to be aware that the expert's judgment may be clouded and makes the expert less likely to be affected by whatever conflict he or she has. This assumption has never been proven and is likely to be incorrect.

Conflicts of interest can exist in all aspects of evaluating evidence. Let's assume that Dr. T and three colleagues submit a review of the early effect of selective serotonin reuptake inhibitors on depression to the *Archives of General Psychiatry*. Before the article is accepted, each author is required to "disclose all potential conflicts of interest, including specific financial interests and relationships and affiliations (other than those affiliations listed on the title page of the manuscript) relevant to the subject of their manuscript" (http://archpsyc.ama-assn.org). Authors are instructed to "err on the side of full disclosure and should contact the editorial office if they have questions or concerns. All such disclosures should be listed in the Acknowledgments section at the end of the manuscript. Authors without conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject of their manuscript, should include a statement of no such interest in the Acknowledgments section of the manuscript." In addition, "authors are expected to provide detailed information about all relevant financial interests and relationships or financial conflicts within the past 5 years and for the foreseeable future (e.g., employment/affiliation; grants or funding; consultancies; honoraria; stock ownership or options; expert testimony; royalties; patents filed, received, or pending; or donation of medical equipment), particularly those present at the time the research was conducted and through publication, as well as other financial interests (such as patent applications in preparation) that represent potential future financial gain" (http://archpsyc.ama-assn.org).

Now suppose that the lead author, a well-known depression researcher, has received honoraria from several pharmaceutical companies and attended meetings sponsored by drug companies. The second author, also a well-known researcher, has received funding for research and consulting, speaker's fees, and travel expenses from a variety of companies that manufacture antidepressant drugs. The third author has received grants from pharmaceutical companies

and from the government. The fourth author is an adviser to several pharmaceutical companies.

Now suppose that the same authors report no financial conflicts of interest. Which publication would one have more confidence in and why? Would you trust a positive outcome from the second group more than from the first and a negative outcome from the first more than from the second? In practice, this information about conflict of interests is not very helpful in judging an article. The first set of authors might have a level of expertise that allows them to interpret the database more accurately than does the second. The second group may have many conflicts but of a different nature. For instance, they may have a bias against the pharmaceutical industry or the use of medications. It is necessary to report such conflicts, but the mere reporting does not really help the reader determine whether the analysis is unbiased. (The example is derived from an actual article: see Taylor et al. 2006.)

The same issues apply to reviewers of articles. A reviewer is just as likely as an author to have conflicts of interest. For instance, reviewers are sometimes selected by looking at the publications at the end of an article. Are reviewers likely to accept papers that support their own ideas or support the "conventional" wisdom and to reject ideas that go against their own ideas or conventional wisdom? Is this bias less serious than that of an investigator who might have some financial gain in a product because he or she holds some shares of the company that produces it? Coyle, the editor of the *Archives of General Psychiatry*, notes: "It would not be reasonable to banish from the ARCHIVES all authors and reviewers who have perceived financial conflicts of interest. This strategy would exclude important clinical research and scientific expertise from these pages" (Coyle and Heckers 2006).

The American Heart Association attempts to address the issue of conflict of interest quite broadly. For instance, members of writing groups are asked to identify "all relationships within the last 2 years that are relevant to the topic of the manuscript." A relationship is "relevant" if the relationship or interest relates to the topic of the manuscript in terms of any of the following: the same or similar subject matter or topic; the same, similar, or competing drug or device, product or service, intellectual property or asset; a drug or device company or its competitor; or has the reasonable potential to result in financial, professional, or other personal gain or loss for you, members of your household, or employer. The author or reviewer needs to answer this in relation to employment, research grants, other research support, speakers' bureau, honoraria, ownership interest, consultation/advisory board, or other. The instructions note that "A relationship is considered to be 'Significant' if (a) the person receives $10,000 or more during any 12 month period, or 5% or more of the person's gross income; or (b) the person owns 5% or more of the voting stock or share of the entity, or owns $10,000 or more of the fair market value of the entity. A relationship is considered to be 'Modest' if it is less than 'Significant' from the preceding definition." Such relationships are important to disclose, but more importantly, the authors or reviewers should be asked to note that their opinions are based on an "unbiased assessment of the facts." In practice, it is often very difficult to evaluate the expertise of the expert.

The "5S" Approach to Searching for Answers

Although many clinicians begin their search with MEDLINE, this is a relatively inefficient strategy because it typically identifies numerous articles that must then be individually reviewed for validity. A more efficient approach is the hierarchical "5S" strategy developed by Haynes (2001a, 2001b, 2006), which involves *systems*, *summaries*, *synopses*, *syntheses*, and *studies* (Table 4–2). At the top of the list are *systems*, by which Haynes means very detailed, often patient-specific data sources, which link the patient's conditions to current best practices. *Summaries* integrate the best available evidence to provide a full range of evidence addressing *all* management options for a given health problem, not just one aspect of the problem as found in single-study resources that make up the lower three categories in the 5S model. *Synopses* are very brief descriptions of original articles and reviews. They can be found under search terms such as *meta-analyses*. *Syntheses* are usually called "systematic reviews" in the research literature. Some databases allow you to limit search results by using that term in the search strategy. Single *studies* are the original journal articles that present the entire report of one study on one aspect or management of a health problem. The distinctions between these categories are somewhat arbi-

TABLE 4–2. The 5S approach to searching for answers

Type of information resource	Examples	Web site
Systems	Electronic records linked to patient guidelines and other evidence	
Summaries	American Psychiatric Association Practice Guidelines	http://www.psychiatryonline.com
	National Guideline Clearinghouse	http://www.guideline.gov
	Clinical Evidence	http://clinicalevidence.bmj.com
	PIER (Physicians' Information and Education Resource)	http://pier.acponline.org/index.html
	National electronic Library for Mental Health	http://www.library.nhs.uk/mentalhealth
	TRIP	http://www.tripdatabase.com
	PubMed	http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed
	Google Scholar	http://scholar.google.com
	Evidence-based textbooks and journals	
Synopses (abstracts/summaries)	*Evidence-Based Mental Health*	http://ebmh.bmj.com
	ACP Journal Club	http://www.acpjc.org
	PubMed	http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed
	Database of Abstracts of Reviews of Effects (DARE)	http://www.crd.york.ac.uk/crdweb
Syntheses (systematic reviews)	Cochrane Database of Systematic Reviews	http://www.cochrane.org/reviews
	DARE	http://www.crd.york.ac.uk/crdweb
Studies	PubMed	http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed
	Google Scholar	http://scholar.google.com
	Cochrane Central Register of Controlled Trials	http://www.mrw.interscience.wiley.com
	National Institutes of Health clinical trials	http://www.clinicaltrials.gov

Source. Adapted from Gray 2004; Haynes 2001a, 2000b, 2006.

trary, and searches in any one category often result in finding articles or outcomes in all of the categories listed in Table 4–2.

In the 5S strategy, priority is given to sources of high-quality, preappraised information so that the clinician can omit the appraisal step (step 3) in the EBM process. Haynes (2006) noted that clinicians practicing in settings with an electronic medical record system may be able to link their patient's characteristics with current evidence-based guidelines for care within that system. If such a system is not available, he recommends that the clinicians then begin to look for summaries. Table 4–2 lists many sources of summaries, discussed later in this chapter.

If no summaries, guidelines, algorithms, or other more comprehensive and systematic reviews are available, Haynes recommends that the clinician then look for brief summaries and synopses over lengthier reviews, assuming that a busy clinician wants an answer as quickly and effortlessly as possible. Although this strategy is appropriate for quickly answering clinical questions in a practice setting, it is not necessarily the most appropriate strategy for performing a comprehensive search of the clinical literature.

Systems

The starting point for the search should be what Haynes (2006) has termed a *system*, an information source that covers a variety of diagnoses, provides a summary of the results of high-quality systematic reviews, and is frequently updated. Such a system would provide the user with a concise summary of the evidence, linked to the original studies. Some clinicians have access to electronic medical record systems with a computerized decision support system that can link a patient's characteristics to current evidence-based guidelines for care. For instance, the Clinical Research Information System (CRIS) used at Duke University, according to their Web site, provides all of the tools to seamlessly allow clinicians to use standard measures to assess patients. This information can be used "to establish and enforce clinical practices and guidelines." Many systems are developing quality assessment metrics that monitor aggregate outcomes related to preestablished standards. In theory, the standards and the practices that lead to them are evidence based.

Summaries (Including Guidelines, Algorithms, and Best Practices)

Many resources are available to help clinicians rapidly find evidence-based information. We have arbitrarily divided these into resources of clinical evidence, clinical guidelines, and algorithms.

Clinical Evidence

The *BMJ* Publishing Group has several publications, products, and activities to promote evidence medicine. *BMJ*'s *Clinical Evidence* is an international peer-reviewed journal that publishes systematic reviews of important clinical conditions. The reviews are frequently updated with the goal of providing busy clinicians with access to the very latest and most relevant medical knowledge for treatment decisions (http://clinicalevidence.bmj.com). *Clinical Evidence* attempts to provide clinically useful information by also including a comments section with recommended options, allowing for interactions among users. *Clinical Evidence* uses Cochrane reviews as a source of high-quality, systematic reviews but also uses reviews of primary studies and reviews of reviews. Reviews are updated annually with a new planned systematic search of databases and include frequent updating through *Clinical Evidence Updates* (www.bmjupdates.com). *BMJ* provides online access, a printed handbook, and a version for a personal digital assistant (PDA). The reviews may cover issues related to an intervention or condition. For instance, the review of smoking cessation ("Putting Evidence Into Practice: Smoking Cessation"; available as a PDF file at: www.clinicalevidence.bmj .com/ceweb/resources/index.jsp) provides information on smoking rates, the disease burden from smoking and health benefits of cessation, predictors of quitting and maintained cessation, public policy on smoking, and interventions.

Another emerging and important source of EBM is PIER (the Physicians' Information and Education Resource). Unfortunately, PIER is only available to American College of Physicians (2009; ACP) members, although some reviews can be accessed online through the *ACP Journal Club* (www.acpjc.org). The *ACP Journal Club* (Haynes 2008) is designed to keep clinicians up to date, but articles may be relevant to an issue a clinician is researching.

The United Kingdom's National Health Service offers an online evidence-based mental health library (NeLMH 2009). This library is collaboratively produced by a core team working for the library with assistance from core groups, including the Centre for Evidence-Based Medicine (www.cebm.net), the Centre for Evidence-Based Mental Health (www.cebmh.com), the Royal College of Psychiatrists, the University of Oxford Department of Psychiatry, and the World Health Organization United Kingdom Collaborating Centre. For instance, the *Schizophrenia Annual Evidence Update 2008* (Mental Health Specialist Library 2008) brings together evidence-based guidelines, systematic reviews, important primary research, service development literature, and patient information.

The TRIP Database (www.tripdatabase.com) began in 1997 with the goal of answering "real" clinical

questions using the principles of EBM. The TRIP Database includes a variety of resources that can be easily searched. The results can be filtered by systematic reviews, evidence-based synopses, guidelines, clinical questions, core-primary research, e-textbooks, and others.

Clinical Practice Guidelines and Algorithms

Table 4–3 lists major guidelines and algorithms available to practitioners. Online access to these may require registration and payment. The most useful starting point in searching for these guidelines is the National Guideline Clearinghouse (www.guideline.gov). This database of clinical practice guidelines is operated by the Agency for Healthcare Research and Quality (AHRQ), in association with the American Medical Association (AMA) and the American Association of Health Plans (AAHP). All guidelines must meet explicit quality criteria for inclusion (Agency for Healthcare Research and Quality 2000). Searching the database yields a structured abstract that is linked to the actual guideline. The database also allows several guidelines to be compared side by side. (General information about clinical practice guidelines and algorithms can be found in Gray and Taylor, Chapter 7 in this volume.)

Most American psychiatrists are familiar with the practice guidelines developed by the American Psychiatric Association (APA). These guidelines cover the major psychiatric disorders and are available both in print form (American Psychiatric Association 2006) and online (www.psychiatryonline.com).

Synopses

If you have no access to a system and found no results from the summaries, the next step is to search for what Haynes (2001a, 2001b, 2006) has termed *synopses*, structured abstracts of high-quality systematic reviews or original articles. The advantages of these resources are that they have already been appraised for quality and are summarized; therefore, busy clinicians can quickly get to the bottom line without reading a lengthy article. Furthermore, these synopses can be accessed through online databases that can be searched quickly. Because these databases contain summaries of only those articles that meet certain quality criteria, they are much smaller than MEDLINE and yield fewer references.

TABLE 4–3. Sources of high-quality clinical practice guidelines and algorithms

Organization	Web site
American Psychiatric Association	http://www.psychiatryonline.com
British Association for Psychopharmacology Consensus Guidelines	http://www.bap.org.uk
Canadian Medical Association	http://www.cma.ca
Clinical Knowledge Summaries (United Kingdom National Health Service)	http://www.cks.library.nhs.uk/home
Expert Consensus Guidelines	http://www.psychguides.com
Harvard South Shore, Psychopharmacology Algorithm Project	http://mhc.com/Algorithms
International Psychopharmacology Algorithm Project (IPAP)	http://www.ipap.org
National Guideline Clearinghouse	http://www.guideline.gov
National Institute for Health and Clinical Excellence (NICE)	http://www.nice.org.uk
New Zealand Guidelines Group	http://www.nzgg.org.nz/index.cfm?
Scottish Intercollegiate Guidelines Network (SIGN)	http://www.sign.ac.uk/guidelines/index.html
Texas Medication Algorithm Project (TMAP)	http://www.dshs.state.tx.us/mhprograms/tmap.shtm
U.S. Department of Veterans Affairs	http://www.va.gov

Evidence-Based Mental Health

For psychiatry, the best source of synopses is *Evidence-Based Mental Health*, published quarterly by the BMJ Publishing Group with the Royal College of Psychiatrists and the British Psychological Society. The online version (http://ebmh.bmjjournals.com) is particularly useful in searching for answers to clinical questions related to common psychiatric disorders.

The staff at *Evidence-Based Mental Health* reviews the major medical and psychiatric journals to identify both original research articles and literature reviews that meet explicit quality criteria. They then prepare structured abstracts, sometimes reanalyzing the data to present it in a uniform format. An outside reviewer supplies a commentary. Finally, a declarative title summarizes the article's "clinical bottom line."

ACP Journal Club

ACP Journal Club, a publication that is similar to *Evidence-Based Mental Health*, is published bimonthly by the American College of Physicians–American Society of Internal Medicine and is available online (www.acpjc.org). Although the intended audience is internists, *ACP Journal Club* does include summaries of articles of psychiatric interest that are related primarily to disorders seen in primary care settings (e.g., depression, anxiety disorders, dementia, delirium, and substance abuse). If you do not have access to *Evidence-Based Mental Health*, then searching the (free) *ACP Journal Club* database may prove useful.

Syntheses

If a relevant synopsis cannot be found, the next step is to search for what Haynes (2001a, 2001b, 2006) has termed a *synthesis*, a high-quality systematic review. A detailed discussion of systematic reviews and how they are appraised for quality is given in Gray, Chapter 6 in this volume. The following discussion concentrates on the methods used for locating systematic reviews. A common method for systematic reviews is to define a protocol at the outset of the study. Greenhalgh and Peacock (2005) found that in a review of diffusion of service-level innovations in health care organizations, only 30% of the sources were obtained from the protocol defined at the outset of the study, 50% were identified by "snowballing" (such as pursuing references of references), and 24% were obtained by personal knowledge or personal contacts. They concluded that systematic reviews of complex evidence cannot rely solely on protocol-driven search strategies; however, "snowballing," personal knowledge, and personal contacts all influence the final database.

Cochrane Database of Systematic Reviews

The single best source of systematic reviews is the Cochrane Database of Systematic Reviews (Glanville and Lefebvre 2000; McKibbon et al. 2002). This database contains high-quality systematic reviews specially prepared by the Cochrane Collaboration, a multisite international workgroup (Antes and Oxman 2001; Cochrane Collaboration 2003). The full-text version of the reviews is available through medical libraries as part of their Ovid database subscription. Abstracts of the reviews are available online without charge (www.cochrane.org/reviews/index .htm).

Database of Abstracts of Reviews of Effectiveness

The Database of Abstracts of Reviews of Effectiveness (DARE) consists of structured abstracts of systematic reviews that meet certain quality criteria (National Health Service Centre for Reviews and Dissemination [NHS CRD] 2000). DARE is maintained by the NHS CRD at the University of York and can be accessed through their Web site (www.crd.york.ac.uk/crdweb). DARE is also available through medical libraries as part of Ovid's Evidence-Based Medicine Reviews, a database that combines Best Evidence, the Cochrane Database of Systematic Reviews, and DARE (Etchells 2000; Ovid 2002b). Because DARE provides only abstracts of reviews, the actual review must be obtained separately if further detail is required.

Health Technology Assessment Database

The Health Technology Assessment database is also maintained by the NHS CRD at the University of York, in collaboration with the International Network of Agencies for Health Technology Assessment (INAHTA; National Health Service Centre for Reviews and Dissemination 2002). It contains abstracts of critical reviews of health technologies, including treatments for psychiatric conditions. Most of the abstracts are linked to the Web site of the agency that produces the report, from which the full text of the document can be obtained. The Health Technology Assessment and DARE data-

bases may be searched simultaneously from the NHS CRD Web site (www.crd.york.ac.uk/crdweb).

Other Sources of Systematic Reviews

Systematic reviews published in journals also may be identified through MEDLINE and similar databases or searched with Google Scholar. This can be accomplished most efficiently with the assistance of "filters" that attempt to limit the search results to systematic reviews or meta-analyses (Glanville and Lefebvre 2000; McKibbon et al. 1999; National Health Service Centre for Reviews and Dissemination 2002; Shojania and Berg 2001). This process is described in more detail in the next section.

Bandolier (www.medicine.ox.ac.uk/bandolier) is an online version of an evidence-based health care journal written by two Oxford scientists. The Web site uses a variety of sources to publish reviews and synopses. If the question being asked happens to fit into an area covered by the service, the information can be useful. For instance, *Bandolier* updated their systematic reviews related to addiction in August 2008. The review was actually a brief synopsis of findings from "three reviews and meta-analyses" of abstinence rates for alcohol addiction for acamprosate, naloxone, and opioid antagonists. The synopsis includes a "clinical bottom line": "Information on interventions for alcohol addictions is sparse. Without any drug intervention, about 1 user in 5 will be abstinent between 3 and 24 months." (Curiously, this conclusion is not supported by the data presented in the rest of the synopses.)

Studies

MEDLINE or similar databases should be used only if the above-mentioned other sources have proved unsuccessful. This is because such a search is apt to yield multiple studies requiring appraisal. In contrast, the resources described in the first three categories are sources of high-quality evidence that have generally already been preappraised for validity, thus sparing the need to go through the critical appraisal step before using the evidence.

Multiple databases can be useful in locating answers to mental health–related questions. However, much of the focus of this section is placed on MEDLINE, partly because psychiatrists are generally most familiar with this resource and can access it most readily.

MEDLINE

MEDLINE is a database maintained by the U.S. National Library of Medicine (NLM). It includes more than 12 million citations, both clinical and preclinical. MEDLINE may be accessed through a variety of different services. Libraries usually obtain the MEDLINE database through a commercial vendor, such as Ovid or SilverPlatter. MEDLINE access is available free of charge through the NLM's PubMed Web site (www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed).

For most psychiatrists, the simplest starting point for a MEDLINE search (using PubMed) is the Clinical Queries interface. The Clinical Queries interface allows one to search by a clinical study category and to limit the scope of the search by category (etiology, diagnosis, therapy, prognosis, or clinical prediction guides) and scope (narrow or broad). PubMed then applies search filters to limit the search to particular types of articles. This filtering is achieved by addition of specific terms to the search, in addition to the terms that the user entered. For instance, Table 4–4 shows how to conduct a search with various filters. In this table, the clinician is interested in articles published since 2002 on the effect of trazodone on sleep but only wants to look at RCTs with a placebo.

TABLE 4–4. Example of a PubMed search

Search: PubMed for the effects of trazodone on sleep
Go to: http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed
Enter: "trazodone and sleep"
Select: Advanced Search
In the field "Search by Author, Journal, Publication Date, and more," type in the Publication Date Search, "2002 to present."
In the field "Limit by Topics, Languages, and Journal Groups," select "Randomized Controlled Trial."
In the "Index of Fields and Field Values," type in "placebo" and select "and."
If you scroll to the top of the page, you will see that you will be searching for randomized controlled trials published since 2002 using the terms (trazodone and sleep) AND (placebo).
The search results in seven studies.

Note. Search performed on June 25, 2009.

The search strategies developed by Haynes et al. (2005) have been embedded in PubMed. For the busy clinician, the Limits function is likely to be more efficient than the Advanced Search mechanism. PubMed now includes several brief tutorials to teach the user about the system.

Other Databases

Although MEDLINE is usually the best starting point for a search, various other specialized databases may be useful in searching for particular types of evidence.

EMBASE, the *Excerpta Medica* Database. EMBASE is an international medical and pharmaceutical database that covers more than 7,000 journals from 70 countries (www.embase.com). EMBASE requires subscribing to the service.

***Cumulative Index to Nursing and Allied Health Literature*.** The Cumulative Index to Nursing and Allied Health Literature (CINAHL) covers 1,000 English-language nursing and allied health journals (www.cinahl.com/library/library.htm). CINAHL requires a subscription.

ClinPSYC and PsycINFO. ClinPSYC and PsycINFO are databases produced by the American Psychological Association. PsycINFO covers 2,150 journals in more than 25 languages. It also includes books, dissertations, and other secondary publications. ClinPSYC is a clinically oriented subset of the PsycINFO database. PsycINFO and ClinPSYC are available through Ovid and many institutions.

Google Scholar. Google Scholar (http://scholar.google.com) is a search engine that searches articles available in the vast Google database. Selecting "Advanced Scholar Search" allows for some search restrictions (e.g., by dates published).

Alternatives to the 5S Approach

The 5S approach by Haynes (2001a, 2001b, 2006) is a stepwise search strategy, beginning with systems and progressing until an answer to the clinical question is found. An alternative strategy is to search several databases simultaneously that include two or more levels in the hierarchy. Two search engines that perform this function are the TRIP Database and SUMSearch.

TRIP Database

The TRIP Database (www.tripdatabase.com) was created in 1997 and is currently the search engine for the National electronic Library for Mental Health. Updated monthly, it is an attempt to link all of the high-quality evidence-based resources available on the Internet. A search that uses the TRIP Database will search the Cochrane Library, DARE, other collections of systematic reviews and guidelines, and even some online journals. It also has links to the PubMed Clinical Queries search interface.

SUMSearch

SUMSearch (http://sumsearch.uthscsa.edu) was developed at the University of Texas Health Science Center at San Antonio. SUMSearch is designed to query only Internet sites that contain evidence written by qualified professionals. Most of the links provided by SUMSearch come from three Internet sites: the NLM, DARE, and the National Guideline Clearinghouse. When querying MEDLINE at the NLM, SUMSearch uses search filters that have been developed by various researchers to search optimally for certain types of articles (Wilczynski et al. 2004). For example, when the clinician clicks the "treatment" focus, SUMSearch includes a search of MEDLINE using a filter validated to find randomized controlled clinical trials. Haase et al. (2007) compared the retrieval efficiency of SUMSearch and Google Scholar. The searches were limited by various terms, "clinical guidelines," and "specific diseases." The performance pattern between search engines was similar: search strategies including the term *guideline* yielded the highest sensitivity (SUMSearch: 81.5%; Google Scholar: 31.9%), and search strategies including the term *practice guideline* yielded the highest specificity (SUMSearch: 89.5%; Google Scholar: 95.7%) and the lowest number needed to read (SUMSearch: 7.0; Google Scholar: 9.3). Thus, SUMSearch was superior to Google Scholar in the retrieval of relevant information (Haase et al. 2007).

Examples of Searching for Evidence With Various Approaches

To explain and compare various approaches, two examples of clinical questions and their resulting search results are given. Of interest, we searched the same question used by Gray in 2004 to determine whether any new findings were available.

Example 1: Cognitive-Behavioral Therapy in Schizophrenia

In patients with schizophrenia, does the addition of cognitive-behavioral therapy (CBT) to usual treatment, compared with usual treatment alone, prevent relapse?

5S Approach

The starting point for this search was a summary, *Clinical Evidence*. We began by searching with the terms *schizophrenia*, *CBT*, and *relapse*. This resulted in a synopsis of a review (Jones et al. 2004), which concluded that CBT compared with standard care does not reduce relapse rates at up to 60 months compared with standard care alone. Evidence was graded as "high-quality." This conclusion differs from the conclusion reached from Gray's search in 2004, which also began with *Clinical Evidence*. In that search, the conclusion was that there is limited evidence from RCTs that CBT may reduce relapse rates. These somewhat contradictory findings point to the need for the practitioner to reassess treatment practices periodically because new evidence may change recommendations.

We then looked at the American Psychiatric Association (2006) "Practice Guideline for the Treatment of Patients With Schizophrenia," which noted that "Overall, the data support the efficacy of cognitive behavior therapy for reducing the frequency and severity of positive symptoms and the distress associated with these symptoms. Furthermore, these gains appear to continue over time. The benefits do not appear to extend to relapse, rehospitalization, or social functioning" (p. 669). With the search terms *schizophrenia* and *CBT*, the NeLMH produced another set of documents, including a 2003 review from the NHSCRD, a 2004 systematic review from the Cochrane database, and the same critical appraisal from the *Evidence-Based Mental Health* database. All three of these searches led to the same conclusion—that CBT is not effective in reducing relapse in patients with schizophrenia compared with usual treatment.

TRIP Database

The terms *cognitive-behavior therapy AND schizophrenia* were searched in the TRIP Database. Several MEDLINE articles were identified, but we then filtered the search by systematic reviews, evidence-based synopses, guidelines, clinical questions, and others. None yielded any useful information. The MEDLINE search, included in the search results, listed some articles that also could have been found by searching MEDLINE directly. We then took a "higher" order approach, first looking for general guidelines, and found that we could search to find specific issues about CBT and schizophrenia. Several articles were returned.

SUMSearch

The terms *cognitive-behavioral therapy AND schizophrenia* were searched in SUMSearch. The search, which took about 90 seconds, produced 112 documents, including one in Wikipedia. (The PubMed Full Text original-research search function did not respond in time; the page suggested searching PubMed directly or trying this search again later.) The search produced three guidelines, including the APA guideline. Under PubMed (possible guidelines), it produced several relevant articles, including a recent review of CBT for medication-resistant schizophrenia (Rathod et al. 2008) and several other guidelines not listed in the TRIP Database. Under PubMed (possible systematic reviews), it also produced some useful articles.

Google Scholar

Google Scholar was searched using the terms *cognitive-behavioral therapy AND schizophrenia*. The search, which took 0.21 seconds, retrieved about 4,400 citations, including reviews and summaries. The material probably included most of the articles listed under SUMSearch.

Example 2: Kava in Generalized Anxiety Disorder

In patients with generalized anxiety disorder, is kava extract more effective than a placebo in relieving symptoms of anxiety?

5S Approach

The search started with a summary (*Clinical Evidence*), but the section on generalized anxiety disorder did not mention kava as a treatment. An online search of *Evidence-Based Mental Health* with the terms *kava AND anxiety* failed to yield any articles. We also looked at clinical guidelines and found no useful information. We then searched the Cochrane

Library, which provided a November 18, 2002, review that concluded the following:

> Systematic literature database. Twenty-two potentially relevant double-blind, placebo-controlled RCTs were identified. Twelve trials met the inclusion criteria. The meta-analysis of seven trials suggests a significant treatment effect for the total score on the Hamilton Anxiety Scale in favor of kava extract. Few adverse events were reported in the reviewed trials, which were all mild, transient and infrequent. These data imply that, compared with placebo, kava extract might be an effective symptomatic treatment for anxiety although, at present, the size of the effect seems to be small. Rigorous trials with large sample sizes are needed to clarify the existing uncertainties. Particularly long-term safety studies of kava are needed.

The DARE database yielded 10 articles but none more recent or relevant than Pittler and Ernst (2003). No articles were retrieved from two other linked databases.

TRIP Database

The terms *kava AND anxiety* were searched in the TRIP Database. Nineteen systematic reviews were found, including Pittler and Ernst (2003). A meta-analysis by Saeed et al. (2007) concluded, "There is substantial evidence that kava has a positive effect on the symptoms of anxiety disorders." This meta-analysis included a Cochrane review, several recent small RCTs and an RCT on safety, and case reports on safety. We then ran a search of the Saeed et al. (2007) article in PubMed to examine "related articles." Doing so can identify articles citing Saeed et al. (2007). One published in 2008 concluded, with regard to herbal treatments, that "kava is effective in reducing anxiety symptoms.... The association of kava with hepatotoxicity, however, is a significant concern" (van der Watt et al. 2008). We had our answer. Kava may be more effective than placebo in relieving symptoms of anxiety, but the risk exceeds the benefit. This search also illustrated that it is worthwhile to look at both benefits and risks.

SUMSearch

The terms *kava AND anxiety* were searched in SUMSearch. The search resulted in two guidelines (but not relevant ones); three PubMed possible guidelines; three DARE articles, one providing information on possible hepatotoxicity; and 22 possible systematic reviews, including those mentioned earlier. SUMSearch also included several articles from Wikipedia that seemed irrelevant to the issue of efficacy and safety. However, we found this ready reference to things such as kava production rather interesting.

Google Scholar

The Google Scholar search produced 3,210 articles in 0.09 seconds. However, it included (at the top) many older and nonrelevant articles.

PubMed

Finally, we ran a search on PubMed. It produced many of the same articles but also some that we had not seen in other searches.

Comparison and Recommendations

After performing several similar searches using different methods, we found that each of the approaches has certain advantages and disadvantages. The 5S approach was most tedious, and much of the material appeared in the other reviews. However, it is a worthwhile place to start for a summary that might be relevant. SUMSearch and the TRIP Database are both slower than a direct search of a database but provide a more comprehensive source of information. Google Scholar is very quick, but PubMed is wonderful in presenting the most recent (but not necessarily most relevant) information first.

On the basis of these admittedly unsystematic and undoubtedly biased observations, we would recommend a somewhat modified 5S approach. We would suggest beginning the search with *Clinical Evidence* because many questions can be answered easily and quickly with this resource. Furthermore, it has the advantage of being available online and in print, and it can be downloaded to a PDA. An alternative is to begin with one's favorite clinical guideline to determine whether information on the topic is available. Reviewing the guideline also allows one to examine the whole treatment approach to a problem.

If online access to *Evidence-Based Mental Health* is available, we recommend searching this database next. It includes synopses of original research that have not yet been incorporated into a systematic review, and the information is presented in a format that will provide an answer quickly.

TABLE 4–5. Suggested search strategy

Step	Resource	Web site
1	*Clinical Evidence*	http://clinicalevidence.com
2	*Evidence-Based Mental Health*	http://ebmh.bmjjournals.com
3	TRIP Database	http://www.tripdatabase.com
4	PubMed Clinical Queries	http://www.ncbi.nlm.nih.gov/corehtml/query/static/clinical.shtml

The third step is to go to clinical guidelines. However, these higher-level approaches may not answer more specific questions. The guidelines had to be reviewed to find statements about kava, for instance.

A quick look at SUMSearch and/or TRIP is worthwhile. The final step in our search would be PubMed. This allows for a very quick perusal of information. Table 4–5 outlines this approach.

Inability to Find an Answer

There are three general reasons for not finding an answer to a clinical question. The first reason has to do with the mechanics of the search process. Changing the search terms, searching by MeSH heading as well as text words, and other techniques described in more detail elsewhere (Greenhalgh 1997, 2001; McKibbon et al. 1999, 2002) may be of value if this is the case.

The second reason has to do with the nature of the question itself. Perhaps the patient population specified in the question is overly specific. Broadening the patient population in the question may result in the search yielding a possible answer. The problem then becomes one of deciding whether that answer can be generalized to a particular patient. This is a topic discussed in more detail in subsequent chapters.

The third reason for not finding an answer is either that the evidence does not exist or that the only evidence is relatively unreliable. The search strategies outlined earlier tend to find evidence that is high in the evidence hierarchy (Table 4–1). If such high-quality evidence does not exist, it is necessary to search for evidence lower in the hierarchy. This can be done through PubMed, by changing the search filters in the Clinical Queries interface from "specificity" to "sensitivity," or by doing a MEDLINE search without filters. Such strategies may result in lower-quality evidence, such as case series or case reports. In some cases, this may represent the best available evidence. Obtaining consultation is another option. Although expert opinion is at the bottom of the evidence hierarchy, it is still better than nothing. Finally, EBM retains a place for clinical judgment (Haynes et al. 2002a, 2002b). In the absence of evidence from the literature, clinical judgment becomes even more important.

References

Agency for Healthcare Research and Quality: The National Guideline Clearinghouse fact sheet [Agency for Healthcare Research and Quality Web site]. July 2000. Available at: http://www.ahcpr.gov/clinic/ngcfact.htm. Accessed June 18, 2002

American College of Physicians: PIER: Physicians' Information and Education Resource. 2009. Available at: http://pier.acponline.org/index.html. Accessed June 25, 2009

American Psychiatric Association: Practice guideline for the treatment of patients with schizophrenia, 2nd edition, in Practice Guidelines for the Treatment of Psychiatric Disorders, Compendium. Washington, DC, American Psychiatric Association, 2006, pp 565–746

American Psychiatric Association: Practice Guidelines for the Treatment of Psychiatric Disorders, Compendium. Washington, DC, American Psychiatric Association, 2006

Antes G, Oxman AD: The Cochrane Collaboration in the 20th century, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 447–458

Antman EM, Lau J, Kupelnick B, et al: A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. JAMA 268:240–248, 1992

Badenoch D, Heneghan C: Evidence-Based Medicine Toolkit. London, BMJ Books, 2002

CINAHL [Edward G. Miner Library Web site]. 2002a. Available at: http://www.urmc.rochester.edu/Miner/Educ/ebnfilt.htm. Accessed June 27, 2003

Clinical Research Information System (CRIS). Available at: http://psychiatry.mc.duke.edu/CMRIS/CMRIndex.htm. Accessed December 2, 2008

Cochrane Collaboration: Introductory leaflets. Available at: http://www.cochrane.org/cochrane/leaflet.htm. May 15, 2003. Accessed June 25, 2008

Coyle JT, Heckers S: Update on the conflict of interest policy for the ARCHIVES. Arch Gen Psychiatry 63:1178, 2006

Davidoff F, Haynes B, Sackett D, et al: Evidence based medicine. BMJ 310:1085–1086, 1995

Etchells E: Ovid. ACP J Club 132:A15, 2000

Glanville J, Lefebvre C: Identifying systematic reviews: key resources. Evid Based Ment Health 3:68–69, 2000

Gray GE: Concise Guide to Evidence-Based Psychiatry. Washington, DC, American Psychiatric Publishing, 2004

Greenhalgh T: How to read a paper: the Medline database. BMJ 315:180–183, 1997

Greenhalgh T: How to Read a Paper: The Basics of Evidence Based Medicine, 2nd Edition. London, BMJ Books, 2001

Greenhalgh T, Peacock R: Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ 331(7524):1064–1065, 2005

Haase A, Follmann M, Skipka G, et al: Developing search strategies for clinical practice guidelines in SUMSearch and Google Scholar and assessing their retrieval performance. BMC Med Res Methodol 7:28, 2007

Haynes RB: Of studies, summaries, synopses, and systems: the "4S" evolution of services for finding current best evidence. Evid Based Ment Health 4:37–39, 2001a

Haynes RB: Of studies, syntheses, synopses, and systems: the "4S" evolution of services for finding current best evidence. ACP J Club 134:A11–A13, 2001b

Haynes RB: Of studies, syntheses, synopses, summaries, and systems: the "5S" evolution of information services for evidence-based healthcare decisions. Evid Based Med 11:162–164, 2006

Haynes RB: ACP Journal Club: the best new evidence for patient care. ACP J Club 148:2, 2008

Haynes RB, Devereaux PJ, Guyatt GH: Clinical expertise in the era of evidence-based medicine and patient choice. ACP J Club 136:A11–A14, 2002a

Haynes RB, Devereaux PJ, Guyatt GH: Physicians' and patients' choices in evidence based practice. BMJ 324:1350, 2002b

Haynes RB, McKibbon KA, Wilczynski NL, et al: Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. BMJ 330(7501):1179, 2005

Jones C, Cormac I, Silveira da Mota Neto JI, et al: Cognitive behaviour therapy for schizophrenia. Cochrane Database Syst Rev 2004 Oct 18;(4):CD000524

Lacchetti C, Guyatt G: Therapy and validity: surprising results of randomized controlled trials, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 13–47

McKibbon A, Eady A, Marks S: PDQ Evidence-Based Principles and Practice. Hamilton, ON, Canada, BC Decker, 1999

McKibbon A, Hunt D, Richardson WS, et al: Finding the evidence, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 247–265

National electronic Library for Health (NeLMH): NHS Evidence—mental health. July 2002. Available at: http://www.library.nhs.uk/MentalHealth. Accessed June 25, 2009

National Health Service Centre for Reviews and Dissemination (DARE): Available at: http://www.crd.york.ac.uk/crdweb. Accessed June 25, 2009

National Library for Health: The Mental Health Specialist Library. 2008. Available at: http://www.library.nhs.uk/mentalhealth. Accessed June 25, 2009

Phillips B, Ball C, Sackett D, et al: Levels of evidence [Centre for Evidence-Based Medicine Web site]. May 2001. Available at: http://www.cebm.net/levels_of_evidence.asp. Accessed June 27, 2003

Pittler MH, Ernst E: Kava extract for treating anxiety. Cochrane Database Syst Rev (1):CD003383, 2003

Rathod S, Kingdon D, Weiden P, et al: Cognitive-behavioral therapy for medication-resistant schizophrenia: a review. J Psychiatr Pract 14:22–33, 2008

Reeves BC, MacLehose RR, Harvey IM, et al: A review of observational, quasi-experimental and randomised study designs for the evaluation of the effectiveness of healthcare interventions, in The Advanced Handbook of Methods in Evidence Based Healthcare. Edited by Stevens A, Abrams K, Brazier J, et al. London, Sage, 2001, pp 116–135

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Saeed SA, Bloch RM, Antonacci DJ: Herbal and dietary supplements for treatment of anxiety disorders. Am Fam Physician 76:549–556, 2007

Shojania KG, Berg LA: Taking advantage of the explosion of systematic reviews: an efficient MEDLINE search strategy. Eff Clin Pract 4:157–162, 2001

Taylor MJ, Freemantle N, Geddes JR, et al: Early onset of selective serotonin reuptake inhibitor antidepressant action: systematic review and meta-analysis. Arch Gen Psychiatry 63:1217–1223, 2006

van der Watt G, Laugharne J, Janca A: Complementary and alternative medicine in the treatment of anxiety and depression. Curr Opin Psychiatry 21:37–42, 2008

Wilczynski NL, Haynes RB, Lavis JN, et al: Optimal search strategies for detecting health services research studies in MEDLINE. CMAJ 171:1179–1185, 2004

This page intentionally left blank

5

Clinical Trials

Gregory E. Gray, M.D., Ph.D.
C. Barr Taylor, M.D.

After identifying one or more articles or other resources to answer a clinical question, the next step in the evidence-based medicine (EBM) process is to appraise the evidence. In this chapter and in Chapters 6, 7, and 9–12 in this volume, guidance on appraising a variety of different types of evidence is provided.

Because questions about therapy are among those most frequently asked by clinicians, we begin with a discussion of individual studies of therapies. In Chapter 6, Gray focuses on the appraisal of systematic reviews of therapies, whereas in Chapter 7, Gray and Taylor focus on the appraisal of treatment guidelines.

Controlled Versus Uncontrolled Trials

Evidence about the effectiveness of a therapy can come from a variety of sources (Table 5–1). However, evidence from some types of studies is apt to be more biased and misleading than evidence from more rigorous study designs (Greenhalgh 2001; Guyatt 2002; Guyatt et al. 2002b; Lacchetti and Guyatt 2002), as explained in this section.

Single Case Reports

The least rigorous and potentially most biased type of evidence is that arising from a single case (Fletcher et al. 1996; Sackett et al. 1991). As clinicians, we like to believe that patients improve as a result of our efforts. We sometimes attribute patient improvement to a therapy when it was, in fact, the result of spontaneous improvement. Published single case reports have the added problem of publication bias: clinicians are more apt to report their successes with a novel treatment than they are their failures (Easterbrook et al. 1991; Fletcher et al. 1996; Montori and Guyatt 2002; Song et al. 2001). Although a single case report might result in a hypothesis that is later tested in a clinical trial, such reports cannot be relied on to guide treatment (Fletcher et al. 1996; Hennekens et al. 1987; Sackett et al. 1991). Single case reports may be more useful, however, in alerting a clinician to possible adverse effects of a treatment, although even here the adverse event may be unrelated to the treatment (i.e., coincidental). In the psychological literature, single case studies (different from reports) have been used to generate hypotheses about potentially useful treatments. One model for a single case design is to use an ABAB approach, in which a baseline is obtained (A), followed by an intervention period (B), a return to baseline (A), and then a reintroduction of the intervention (B).

Case Series

A case series is only slightly better than a single case report (Fletcher et al. 1996). Here the author reports on a series of patients treated for a particular condition. Because patients are not enrolled in a formal study, it is difficult to know whether the results reflect all patients receiving a particular therapy or whether they reflect only the successes. As with single case reports, publication bias is a significant

TABLE 5–1. Types of studies used to address treatment effectiveness

Uncontrolled studies
Single case reports
Case series
All-or-none case series[a]
Uncontrolled clinical trials
Controlled studies
Studies with historical controls
Studies with concurrent nonrandomized controls
Patients of other physicians or clinical sites
Patient or physician choice of treatment
Systematic allocation
Randomized controlled trials[a]
With blinding[a]
Without blinding (open-label study)

[a]Strongest study designs.

problem in case series, and they are best regarded as a source of ideas for further study.

A particular type of case series, however, does rank high in the evidence hierarchy: the "all-or-none" case series (Badenoch and Heneghan 2002) (Table 4–1, Chapter 4). An all-or-none case series refers to a series of patients who receive treatment for a disease with a universally bad (usually fatal) outcome. If a new treatment leads to a better outcome in all of the patients treated, there is strong evidence of treatment effectiveness. Examples of such an all-or-none case series are some of the earliest reports of the effectiveness of antibiotics in treating infections, which were performed 6 decades ago. It is difficult, however, to think of a psychiatric disorder for which an all-or-none case series would be an appropriate study design.

Uncontrolled Clinical Trials

An uncontrolled clinical trial is a step up from a case series. In this study design, patients enrolled in the study receive the new treatment, but there is no control group. Unlike in a case series, there is a study protocol in an uncontrolled clinical trial that specifies the nature of the subjects, treatment, and outcome measures. Because no control group is used, various factors can bias the results. These factors include the Hawthorne, experimenter expectancy (Pygmalion), and placebo effects; observer bias; regression toward the mean (Yudkin and Stratton 1996); and the natural history of the illness (Fletcher et al. 1996).

Historical Controls

Because improvement can occur spontaneously, as well as being a result of treatment, more convincing evidence about the effectiveness of a therapy comes from controlled trials. However, various control groups are sometimes used in clinical trials (Altman 1991; Bland 2000; Fletcher et al. 1996). The first type is a historical control. In this situation, the outcomes of the patients receiving the experimental therapy in the trial are compared with the outcomes of the patients with the same disorder and in the same treatment setting but who have received another therapy in the past. An example is the comparison of the lengths of stay of inpatients with schizophrenia treated with atypical antipsychotic medications with those of inpatients with schizophrenia treated in the same institution a decade earlier, prior to the introduction of atypical antipsychotic medications. Although the use of historical controls might be convenient, the results can be biased by changes in diagnostic criteria, patient acuity and demographics, and other aspects of care that may have occurred over time (Fletcher et al. 1996; Hennekens et al. 1987). Furthermore, information about the historical control group typically comes from medical records recorded for purposes other than research, whereas the information about patients receiving the experimental therapy is collected with the purpose of the study in mind. This difference in the type of information collected is a form of observer bias (Daly and Bourke 2000). As a result, clinical trials using historical controls often may provide misleading answers and generally overestimate the true treatment effectiveness (Altman 1991; Bland 2000; Everitt 1989; Fletcher et al. 1996).

Nonrandomized Clinical Trials

In other types of nonrandomized clinical trials, subjects are nonrandomly assigned to one or more therapies (Altman 1991; Bland 2000; Fletcher et al. 1996). Examples include studies comparing treatments at one clinic with treatments at another clinic

or treatments by one set of providers with treatments by another set of providers. Other examples include studies in which patients volunteer for the treatment they are to receive or in which they are systematically allocated to a treatment group (e.g., every other patient is assigned to a particular treatment). However, the key characteristic of such studies is that the subjects are not randomly assigned to the clinic, provider, or therapy. As a result, the various groups of subjects may differ at the outset of the study in ways that affect their prognosis, which is referred to as "selection bias" or "allocation bias" (Daly and Bourke 2000; Fletcher et al. 1996). Such studies are not true experiments but are instead a type of observational study known as a *cohort study*.

Randomized Controlled Trials

The preferred study design to assess the effectiveness of a therapy is a randomized controlled trial (RCT) (Altman et al. 2001). In this study design, subjects are randomly assigned to either the control treatment or one or more experimental treatments at the onset of the trial. Such randomization serves to make the control and experimental groups comparable in characteristics that may influence prognosis (confounding factors) (Altman 1991; Altman et al. 2001; Daly and Bourke 2000; Fletcher et al. 1996). In addition, conventional tests of statistical significance are all based on the assumption of random assignment of subjects (Altman 1991; Cummings et al. 2001; Daly and Bourke 2000). The blinding of subjects, clinicians, and raters is an additional attempt to reduce sources of bias in RCTs (see next section).

Controlled usually refers to comparing an active treatment with no treatment or a placebo treatment. However, if one treatment is considered the standard, a common type of pharmaceutical study is to compare the active treatment with the new treatment, with the assumption that the new treatment needs to perform as well as the older treatment. In this model, the investigators may be satisfied if the new treatment performs "as well" as the older treatment. Such studies usually need to have a very large number of subjects to be meaningful; that is, to have sufficient power to detect small differences. Otherwise it may appear that the new and old treatments are the same, but mainly because the study did not have adequate power to detect differences, leading to a type II error.

Sources of Bias

The clinician interested in a particular therapy wants an unbiased estimate of how it compares with another therapy or placebo. The results of a particular study may over- or underestimate the true difference in effectiveness of the two therapies. One reason for this is chance (random error), discussed further in the next section. The other reason is bias (defined as a systematic deviation from the true results), which results in either a systematic overestimation or a systematic underestimation of treatment effectiveness (Guyatt 2002; Sitthi-amorn and Poshyachinda 1993). One goal, therefore, in study design is to minimize bias (Altman 1991).

Confounding

One source of bias in a clinical trial is having experimental and control groups that differ at the onset of the study in characteristics that affect outcome (Guyatt 2002). This is a form of selection bias, and the subject characteristics that affect outcome are known as *confounding factors* (Altman et al. 2001; Daly and Bourke 2000; Jadad 1998). Confounding factors can include variables such as age, sex, ethnicity, illness severity, and comorbid illnesses. Adjustment for known confounding factors can occur in the statistical analysis of a study (Daly and Bourke 2000). Such statistical techniques cannot, however, adjust for unknown confounding factors. Randomization will, however, automatically adjust for such confounding factors by tending to make the treatment and control groups similar (Altman 1991; Altman et al. 2001; Daly and Bourke 2000; Fletcher et al. 1996; Guyatt 2002). Even with randomization, there will still be some differences between the experimental and control groups, but the statistical techniques that are used to analyze studies assume a certain amount of chance variation and take it into account (Daly and Bourke 2000). Indeed, if some form of matching is used to attempt to further minimize differences in known confounding factors between the experimental and control groups, this must be taken into account in the statistical analysis; if it is not, tests of statistical significance will be overly conservative (Daly and Bourke 2000; Peto et al. 1976).

Hawthorne Effect

Several nonspecific effects also can bias the results, including Hawthorne, Pygmalion, and placebo ef-

fects. The Hawthorne effect was first observed in studies of worker productivity at the Hawthorne Western Electric plant in Illinois in 1924 and refers to the tendency of subjects to do better solely because they are being studied (Fletcher et al. 1996; Holden 2001). Some of this may involve subject expectations, but it can also be the result of nonspecific effects of the study situation, such as the increased attention received. The Hawthorne effect is one reason that studies involving historical controls will produce biased results: experimental subjects know that they are part of a study, and they therefore show the Hawthorne effect, whereas historical control subjects were not originally experimental subjects and hence did not show such an effect. The Hawthorne effect affects both the experimental and the control groups in an RCT and thus is eliminated as a source of bias in this study design.

Pygmalion Effect and Co-interventions

The Pygmalion effect, also called the *experimenter expectancy effect*, was first described in educational research, in which it was found that teacher expectations affect pupil performance (Rosenthal and Jacobson 1968). Subsequent research has indicated that experimenter expectations regarding treatment effect may result in differential attention or interactions with some subjects, which results in a change in subject behavior in the direction of experimenter hypothesis (Rosenthal and Rosnow 1991). A related effect is that of a clinician involved in a trial providing additional care (e.g., time, support) to patients in one treatment group but not to those in the other treatment group. This effect is known as *co-intervention* or *performance bias* (Altman et al. 2001; Fletcher et al. 1996; Guyatt 2002; Jüni et al. 2001). In clinical trials, this effect can be minimized by blinding clinicians to the treatment being provided. Such blinding can occur in drug trials; however, such blinding is obviously impossible in trials of psychotherapy. It can, however, be minimized by documenting adherence to treatment protocols (Guyatt 2002).

Placebo Effect

The third nonspecific effect is the placebo effect (Crow et al. 2001; Kaptchuk 1998; Laporte and Figueras 1994). In this effect, it is the subject's expectation of improvement, combined with other nonspecific psychotherapeutic effects, that leads to improvement (Chaput de Saintonge and Herxheimer 1994; Crow et al. 2001). The magnitude of the placebo response rate varies by disorder; it is greater in depression and anxiety disorders than in schizophrenia, but even in acute mania, there is a sizable placebo response (Charney et al. 2002) (Table 5–2). To separate the specific effects of a therapy from the nonspecific placebo effects, it is necessary to have a control group. Blinding the patient to the therapy being administered further reduces the placebo effect (Altman et al. 2001; Guyatt 2002), although this is far more difficult with psychosocial interventions than with drug therapies.

Observer, Detection, or Ascertainment Bias

The final important source of bias in a clinical trial is observer, detection, or ascertainment bias (Altman et al. 2001; Jadad 1998; Jüni et al. 2001). If an interviewer or rater knows which treatment a patient is receiving, he or she may differentially inquire about certain symptoms or see improvement where none exists. Blinding the interviewer or rater to the treatment is an important way of minimizing observer bias (Altman et al. 2001; Guyatt 2002; Jüni et al. 2001); however, it is difficult to do this entirely because the treatment received sometimes may be discerned from treatment side effects or from a patient's comments.

Minimizing Bias Through Blinding

To minimize sources of bias, the optimal study design is an RCT in which the subjects, clinicians, and raters are all blind to the treatment being administered. This is possible in drug trials; however, in many trials involving psychosocial interventions, only the rater can be blinded. The terms *single blind*, *double blind*, and *triple blind* are often used to describe the design type of a study, but there is little agreement about the meaning of the terms (Devereaux et al. 2001; Montori et al. 2002). Consequently, it is preferable to simply specify which of the various participants in a study are blinded to the treatment (Altman et al. 2001; Devereaux et al. 2002; Fletcher et al. 1996).

Basic Statistical Concepts

In reviewing the results of a clinical trial, it is important to have at least a basic understanding of biosta-

TABLE 5–2. Placebo response rates in psychiatric disorders

Disorder	Outcome measure	Study duration	Placebo response rate (%)
Schizophrenia, acute episode	40% reduction in BPRS	6 weeks	8–32
Schizophrenia, maintenance	No relapse	9 months	34
Bipolar disorder, acute mania	50% reduction in YMRS	3 weeks	24
Bipolar disorder, maintenance	No relapse	2 years	19
Major depression	50% reduction in Ham-D	4–24 weeks (average 6 weeks)	30
Panic disorder	50% decrease in attacks	12 weeks	50
Social phobia	Much or very much improved	8–14 weeks	17–32
Obsessive-compulsive disorder	35% reduction in Y-BOCS	9–13 weeks	8–60

Note. BPRS=Brief Psychiatric Rating Scale; Ham-D=Hamilton Rating Scale for Depression; Y-BOCS=Yale-Brown Obsessive Compulsive Scale; YMRS=Young Mania Rating Scale.
Source. Data from Cookson et al. 2002; Walsh et al. 2002; Woods et al. 2001.

tistics. The following qualitative discussion does not discuss specific methods of statistical analysis or calculations; for such topics, the reader is referred to several excellent texts (Altman 1991; Bland 2000; Daly and Bourke 2000).

Hypothesis Testing

We begin with a hypothetical clinical trial comparing an experimental therapy with a control therapy. In this trial, participants are randomly assigned to one of two treatment groups. At the end of the trial, we assess the outcome in the two groups and find a difference. How do we know whether this difference is because the treatments differ in effectiveness or whether it is the result of chance? We know that even if the two treatments are identical, by chance alone there could be some difference in outcome. We therefore need to set a threshold, with differences in outcome greater than that threshold unlikely to be the result of chance alone. Conventionally, this threshold is set so that there is a 5% chance that a difference of that magnitude (or greater) will be the result of chance alone. In practice, an appropriate test statistic (e.g., t, F, or χ^2) is computed, from which a P value is derived. If $P<0.05$, the difference is considered *statistically significant*. This approach to data analysis is referred to as *hypothesis testing*, in that a difference in treatment effect that exceeds a threshold leads us to reject the null hypothesis that there is no difference between treatments.

The results of this clinical trial also can be conceptualized as falling into one of three categories: uncontrolled studies, controlled but not randomized studies, and randomized controlled trials (Table 5–3). In reality, the experimental and control treatments are either equivalent or different. In our experiment, either the difference exceeds the preset threshold and is considered "statistically significant" or it does not. Let us now consider some of the various combinations.

Type I Errors

If there is truly no difference between the experimental and the control treatments, our experiment, by chance alone, might find a difference large enough to be called "statistically significant." This is the equivalent of a *false-positive* result on a diagnostic test. In statistical terms, it is considered a type I error. In the example in the previous subsection, the threshold has been set so that a type I error occurs less than 5% of the time. In statistical terms, it is represented as $\alpha=0.05$, where α is the probability of a type I error.

Type II Errors

Now suppose that a difference truly exists between the two treatments. In our experiment, sometimes we find a large difference between the treatment groups, and sometimes, by chance alone, we may find only a small difference. Because we have set a threshold such that only differences that exceed the

TABLE 5–3. Possible outcomes of a clinical trial

	True difference between treatments	
Study outcome	No true difference	True difference exists
Difference found	False-positive result Type I error Probability=α	True-positive result Power=1 – β
No difference found	True-negative result	False-negative result Type II error Probability=β

threshold are considered *statistically significant*, some of the results may not be considered *significant*, even though there is truly an underlying difference in the effectiveness of the two treatments. This is considered a type II error and is the equivalent of a *false-negative* result from a diagnostic test. The probability of a type II error occurring is represented by β.

Power

Ideally, we would like to have a high probability of detecting a difference when such a difference truly exists. This is the equivalent of a *true-positive* test result. Such a probability is given the term *power* and is represented by 1 minus β.

The magnitude of the treatment effect needed for it to be considered statistically significant is determined by the variability of the results and by the number of subjects studied. The more variable the data, the larger the difference that must be obtained; the larger the number of subjects, the smaller the treatment difference needed for statistical significance. The same is true regarding the power of the experiment to detect a real treatment effect: the more variable the data, the larger the number of subjects required. In general, small studies often lack the power to detect clinically significant differences in treatment effectiveness; for that reason, such studies are considered by some to be unethical (Collins and MacMahon 2001; Halpern et al. 2002).

Confidence Intervals

The approach to data analysis that has been summarized thus far is the classical approach of *hypothesis testing*, in which a difference in treatment effect that exceeds a threshold is said to reject the null hypothesis that there is no difference between treatments. Such an approach focuses on *P* values and statistical significance, but it largely ignores the magnitude of any difference found (Sterne and Smith 2001).

An alternative approach that has become popular in recent years involves *confidence intervals* (CIs) (Gardner and Altman 2000). In this approach, the difference in treatment effectiveness between groups in the clinical trial is used to construct a CI. If a 95% CI is constructed, it implies that there is a 95% chance that the true difference in effectiveness lies within that interval. An interval that does not include zero is the equivalent of rejecting the null hypothesis (i.e., the equivalent of a statistically significant result).

There are several advantages to using CIs instead of hypothesis testing (Gardner and Altman 2000; Guyatt et al. 2002c). First, CIs provide a range in which the true treatment effectiveness is expected, with a narrow CI implying a precise estimate of treatment effectiveness. Second, in negative studies in which the null hypothesis is not rejected, CIs may suggest that a clinically important difference is present but that the power of the study to detect it was too low. There is a difference, for example, between a wide CI that barely overlaps zero and a narrow CI that centers on zero. In the first case, the width suggests that the study was too small to provide a precise estimate and that a large treatment difference cannot be excluded. In the second case, the estimate is quite precise and implies that any difference that exists is too small to be clinically important. Finally, CIs are useful in systematic reviews and meta-analyses (see Gray, Chapter 6 in this volume).

Measures of Treatment Effectiveness

There are many different approaches for describing the effectiveness of a treatment (Jaeschke et al. 2002; Sackett et al. 1991). A drug company promoting a new medication may choose the measure that puts the drug in the best light. However, for a clinician choosing a therapy or advising a patient, there is a need for measures that accurately reflect how one treatment compares with either a placebo or another active treatment.

In psychiatric research, especially drug trials, investigators frequently use a variety of rating scales. In reporting the results of a study, they may compare the differences in rating scale scores of patients in the experimental and control groups. Although the use of such rating scales may be necessary for U.S. Food and Drug Administration (FDA) approval and may generate data that can be easily analyzed, they also have limitations. For example, clinicians and patients cannot easily appreciate the practical implications of a small difference on a particular scale.

In contrast, dichotomous outcome measures are more clinically useful than rating scales. Examples of such outcome measures include dying, being readmitted to the hospital, achieving full remission, being rated as "improved" or "much improved," or having at least a 50% decrease in score on a rating scale. These are all measures that most clinicians and patients consider clinically significant and that can be more readily understood than continuous rating scales that yield a numerical score. In addition, the use of such dichotomous outcome measures allows for the calculation of several useful measures of clinical importance. In illustrating these various measures and their calculation, we have used the results of a hypothetical RCT, comparing an antidepressant drug with a placebo for the treatment of major depression; the outcome measure was "remission" (see Table 5–4).

Percentage Response

In our hypothetical experiment, 60% of the patients receiving the antidepressant drug responded. This is the simplest way of expressing the effectiveness of the antidepressant drug; however, it fails to take into account the high percentage of patients who responded to the placebo. To do so, a better measure of treatment response is thus necessary.

TABLE 5–4. Results of hypothetical experiment used to illustrate measures of treatment effect size

Treatment	Not responding (%)	Responding (%)
Placebo	60	40
Antidepressant	40	60

Control event rate (CER)=60%
Experimental event rate (EER)=40%
Relative risk=EER/CER=40/60=0.67
Relative risk reduction=(CER–EER)/CER=20/60=0.33
Odds ratio=[EER/(1–EER)]/[CER(1–CER)]=[(40/60)/(60/40)]=0.45
Absolute risk reduction (ARR)=CER–EER=60%–40%=20%
Number needed to treat=1/ARR=1/(0.2)=5

Relative Risk

In our example, 40% of the patients receiving the placebo responded and 60% of the patients receiving the placebo did not respond. By convention, when relative risk (RR) is used as a measure of treatment effectiveness, results are expressed in terms of a bad outcome (e.g., nonresponse). An effective treatment is therefore one that reduces bad outcomes. With this convention, the percentage of control subjects *not* responding is called the *control event rate* (CER), which in this example is 60%.

For those receiving the antidepressant, 60% of the patients responded, whereas 40% of the patients did not respond. The percentage of experimental subjects *not* responding is called the *experimental event rate* (EER), which in this example is 40%.

Relative risk is the ratio of EER to CER. In this case, RR=0.4/0.6=0.67, which means that patients receiving medication had only two-thirds the nonresponse rate of the placebo group. More effective treatments provide greater reductions in the risk of a negative outcome. RR values for effective treatments vary between 0 and 1, with smaller values indicating more effective treatments.

RR provides a comparison of the experimental and control treatments, but it can be misleading. For example, reducing the nonresponse rate from 90% to 45% yields the same RR as reducing it from 2% to 1%. In the first example, the treatment causes a much greater percentage of patients to respond than

in the second example; however, both examples yield the same RR.

Relative Risk Reduction

Relative risk reduction (RRR) is calculated with the formulas (CER–EER)/CER or RRR=1–RR. Using the data in Table 5–4, RRR=0.2/0.6=0.33, which means that the nonresponse rate was decreased by one-third. Like RR, RRR varies between 0 and 1 for effective treatments; however, in this case, larger values (i.e., values closer to 1) indicate a more effective treatment. RRR has the same limitations as RR.

Odds Ratio

The odds ratio (OR) is a measure of treatment effect that is similar to RR. The odds of an event occurring are expressed as the ratio of the probability of the event occurring to the probability of the event not occurring. In our example, the odds of nonresponse are 1.5 (or 60 to 40) in the placebo group and 0.67 (or 40 to 60) in an experimental group. The OR is simply the ratio of the odds of a bad outcome in the experimental group divided by the odds of a bad outcome in a control group: 0.67/1.5=0.45. As with RR, for effective treatments, the OR varies between 0 and 1, smaller values being associated with a greater treatment effect. The OR is often used as an effect measure in meta-analyses because of its statistical properties (Deeks and Altman 2001).

Absolute Risk Reduction

Absolute risk reduction (ARR) is simply the difference between the CER and the EER. Using the data from Table 5–4, ARR=60% –40%=20%, which means that 20% fewer patients taking medication failed to respond; conversely, 20% more patients taking medication responded. Because this is an absolute (not relative) measure, it can be used to estimate the percentage of patients undergoing treatment who will benefit more from the experimental treatment than from the control treatment. ARR for effective treatments varies from 0% to 100% (or 0 and 1, if not expressed as a percentage), with larger values indicating more effective treatments. Unlike RR and RRR, ARR provides a measure of how many patients receiving treatment will benefit from it and thus avoids some of the limitations of these prior measures.

Number Needed to Treat

The measure believed by many to be the best expression of relative treatment effectiveness is number needed to treat (NNT) (Cook and Sackett 1995; Jaeschke et al. 2002; Laupacis et al. 1988; Sackett et al. 1991, 2000; Szatmari 1998). NNT is simply the reciprocal of the ARR. In our example, NNT= 1/0.2=5, which means that for every five patients treated with medication, there will be one less case of nonresponse than if all patients had received the control treatment. Put another way, for every five patients treated with medication, there will be one additional patient who responds to medication who would not have responded to the placebo. The calculation of CIs for NNT is given by Altman (1998) and can be found in Appendix B.

NNT is believed by most clinical epidemiologists to be the least misleading and the most clinically useful measure of treatment effectiveness, although patients sometimes have difficulty understanding the concept (Kristiansen et al. 2002). Inclusion of NNT in the reporting of the results of clinical trials is recommended in the Consolidated Standards of Reporting Trials (CONSORT) guidelines (Altman et al. 2001; Begg et al. 1996) (statement available at www.consort-statement.org). Despite this recommendation, relatively few trials report NNTs (Nuovo et al. 2002), so it is often necessary for the reader to go through the calculations himself or herself with the data presented in a published article.

Examples of NNTs for common therapies for psychiatric disorders are given in Table 5–5. As can be seen, most psychiatric therapies have NNTs in the range of 3 to 6, which means that for every three to six patients treated, there is one good outcome that would not otherwise have occurred. For comparative purposes, in a 5-week follow-up of patients with acute myocardial infusion, using death as an outcome measure, streptokinase infusion had an NNT of 15; in a 5.5-year follow-up of patients with moderate hypertension (diastolic blood pressures of 90–109 mm Hg), using death, stroke, and myocardial infarction as outcomes, antihypertensive drugs had an NNT of 128 (Sackett et al. 2000). Hence psychotropic medications are relatively effective when compared with other classes of drugs used in medicine. As another example, Citrome (2008) used NNT to compare new treatments for depression, schizophrenia, and bipolar disorder.

TABLE 5–5. Examples of number needed to treat (NNT) for common psychiatric disorders and treatments

Disorder	Treatment comparison	Outcome measure	NNT
Major depression	Antidepressant vs. placebo	50% reduction in Ham-D	3
	IPT vs. clinical management	Recovery	5
	CBT plus antidepressant vs. monotherapy	50% reduction in Ham-D	5
Acute mania	Valproate or lithium vs. placebo	50% reduction in SADS-M	5
Bipolar disorder	Lithium vs. placebo	Relapse	3
Schizophrenia	Antipsychotic vs. placebo	40% reduction in BPRS or "much improved" CGI Scale	2–5
	Family intervention vs. usual care	Relapse	7
Panic disorder	SSRI vs. placebo	Panic free	3–6
Social phobia	Paroxetine vs. placebo	"Much improved" CGI Scale	3
	Group CBT vs. placebo	"Much improved" CGI Scale	3
Obsessive-compulsive disorder	SSRI vs. placebo	35% reduction in Y-BOCS	4–5
Bulimia nervosa	Antidepressant vs. placebo	Remission	9

Note. BPRS=Brief Psychiatric Rating Scale; CBT=cognitive-behavioral therapy; CGI=Clinical Global Impressions; Ham-D=Hamilton Rating Scale for Depression; IPT=interpersonal psychotherapy; SADS-M=Schedule for Affective Disorders and Schizophrenia, mania component; SSRI=selective serotonin reuptake inhibitor; Y-BOCS=Yale-Brown Obsessive Compulsive Scale.
Source. Data from Cookson et al. 2002; Geddes and Butler 2001; Hay and Bacaltchuk 2001; McIntosh and Lawrie 2001.

Critical Appraisal Guide for Therapy Studies

In this section of the chapter, we introduce a structured approach to the critical appraisal of therapy studies, which is mirrored in subsequent chapters that deal with other types of studies. The guidelines for appraisal of therapy studies are found in Table 5–6.

Is the Study Valid?

Before considering the results of a study, one must first focus on its "Methods" section to assess the study's validity. The first question to ask is whether the study was randomized with a concealed randomization list. Randomization minimizes bias that might result from patients with different prognoses being enrolled in either the experimental or the control treatment group (Altman et al. 2001; Collins and MacMahon 2001; Fletcher et al. 1996; Jüni et al. 2001; Lacchetti and Guyatt 2002; Sackett et al. 1991, 2000). Studies comparing randomized and nonrandomized trials of the same therapies have noted several instances of therapies that seemed effective in nonrandomized trials being much less effective or ineffective in randomized studies, although this is not always the case (Ioannidis et al. 2001; Kunz and Oxman 1998; Lacchetti and Guyatt 2002). Of equal importance, however, is that the allocation list is concealed. There have been some instances in which investigators have been able to determine the assignment of the next patient to be enrolled and have used that information to systematically enroll sicker patients into one treatment group (Altman and Schulz 2001). In addition, studies comparing trials in which allocation was concealed with those in which it was not concealed have found important differences in the size of the treatment effect (Altman and Schulz 2001; Jüni et al. 2001; Kunz and Oxman 1998; Lacchetti and Guyatt 2002; Schulz 2000).

The next question to ask is whether subjects and clinicians were blinded to the treatment that was administered. As noted earlier, blinding represents an attempt to prevent observer bias, placebo effects, and experimenter expectancy effects from being responsible for any observed difference between the

TABLE 5–6. Critical appraisal guide for therapy studies

Is the study valid?
Is it a randomized controlled trial?
Was the randomization list concealed?
Were subjects and clinicians blinded to treatment being administered?
Were all subjects enrolled in the trial accounted for?
Were subjects analyzed in the groups to which they were assigned?
Despite randomization, were there clinically important differences between groups at the start of the trial?
Aside from the experimental treatment, were the groups treated equally?
Are the results important?
How large was the treatment effect (e.g., the number needed to treat)?
How precise were the results (e.g., the width of confidence intervals)?
Can the results be applied to my patient?
Is my patient too different from those in the study?
Is the treatment consistent with my patient's values and preferences?
Is the treatment feasible in this setting?

Source. Adapted from Gray 2002, 2004; Guyatt et al. 2002a; Sackett et al. 2000.

experimental and the control groups. In general, nonblinded studies overestimate the true treatment effect size (Jüni et al. 2001; Schulz 2000). Third, one should check whether all subjects who entered the trial were accounted for at its conclusion and whether they were analyzed in the groups to which they had been randomly assigned.

In 1993, 30 experts, including medical journal editors, clinical trialists, epidemiologists, and methodologists, met with the aim of developing a new scale to assess the quality of RCT reports. However, during preliminary discussions, participants thought that many of the suggested scale items were irrelevant because they were not regularly reported by authors. In fact, accumulating evidence indicated that the quality of reports of RCTs was unsatisfactory (Altman and Doré 1990; Pocock et al. 1987). Therefore, the group began to focus on ways to improve the reporting of RCTs, which resulted in recommendations for the standardized reporting of clinical trials (Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature 1994). In 2001, the group created the CONSORT Statement (www.consort-statement.org) to help authors improve reporting of two parallel-design RCTs by using a checklist and flow diagram (Altman et al. 2001).

Accounting for all subjects is made easier when a flow diagram of subject progress through the phases of a randomized trial is included in the article, as suggested by the CONSORT guidelines (Altman et al. 2001; Moher et al. 2001). This is important because substantial loss of subjects to follow-up can seriously bias the results, an effect referred to as *attrition bias* (Guyatt and Devereaux 2002; Jüni et al. 2001). It is sometimes stated that at least 80% follow-up is sufficient for the results to be valid (Streiner and Geddes 2001), although other authors argue that this rule of thumb is misleading (Guyatt et al. 2002a). Intention-to-treat analysis, which includes data on patients who did not complete the trial in their assigned group, is the preferred method of analyzing the data, although the method has limitations (Altman et al. 2001; Collins and MacMahon 2001; Guyatt and Devereaux 2002; Guyatt et al. 2002a; Jüni et al. 2001; Streiner and Geddes 2001).

Many journals now require clinical trials to include a CONSORT flowchart (available at www.consort-statement.org/index.aspx?o=1031) and to provide evidence that issues raised in the CONSORT statement have been followed.

Many clinical trials are now available through www.clinicaltrials.gov, which includes more than 62,000 trials from 157 countries. To be registered, the

trial sponsor needs to complete a detailed description of the study. The interested reader can find how to register a trial at http://prsinfo.clinicaltrials.gov. Trial registries acceptable to the *Archives of General Psychiatry* include www.clinicaltrials.gov, http://isrctn.org, http://actr.org.au, http://trialregister.nl, and www.umin.ac.jp/ctr.

Finally, clinicians should determine whether other differences between the control and the experimental groups could bias the results. For example: Despite randomization, were there significant differences between the two groups at the start of the trial that could have affected the outcome (i.e., confounding)? Aside from the experimental treatment, were the groups treated equally (i.e., no co-intervention)?

Only after the validity has been appraised should one turn to the results. After all, if the results are not valid, it does not matter what they show.

Are the Results Important?

When reviewing the results of a study, clinicians should give some thought as to what would be considered a clinically significant outcome measurement. As noted earlier, these are often dichotomous outcome measures, and NNT should be calculated if it is not given in the results.

Can the Results Be Applied to Other Patients?

The final step in the appraisal process is to assess whether the study results can be applied to other patients. Several considerations enter into this assessment. First, is the patient so different from the patients in the studies that the results do not apply? The question here can be reframed as: "Is the pathobiology of this patient so different from that of the study patients that the results cannot apply?" In general, the answer to this question is "no" (Sackett et al. 1991, 2000; Straus and McAlister 2001). However, quantitative differences may apply. In particular, if a patient's likelihood of improving *without* the experimental treatment is greater than that of the patients in the study, then the patient will receive less benefit from the treatment than expected, according to the calculated NNT (Sackett et al. 2000). Methods of calculating NNT related to outcomes can be found in Appendix B. Furukawa et al. (2002) confirmed the validity of this approach.

Two additional questions must be asked before the clinician applies the treatment to his or her patient. The first is whether the treatment is consistent with the patient's values and preferences. The second is whether the treatment is feasible in the clinician's setting. Assuming that the clinician has identified a valid study which has found an effective treatment that can be applied to his or her patient, the next steps in the EBM process (Table 2–1, Chapter 2) are to apply the treatment and then to evaluate the outcome.

Industry-Sponsored Trials

Industry-sponsored trials are a major source of information about the effectiveness of new medications. A recent review of 397 articles published over 3 years in four psychiatric journals found that 60% were industry supported (Perlis et al. 2005). Astonishingly, a systematic review of drug treatments of insomnia by the Agency for Healthcare Research and Quality found that all but 5 of 56 RCTs were apparently funded by the industry (Buscemi et al. 2007). Industry-sponsored trials accounted for most of the results used in the meta-analysis, and results from the meta-analysis were used to make evidence-based treatment recommendations.

From a marketing standpoint, pharmaceutical companies are interested in showing that their product is equally as effective as competing medications or may have advantages, such as fewer side effects (M. Angell 2004; Safer 2002). New medications also must be shown to be safe. The long-standing concerns about the role of industry in sponsoring trials and presenting biased results were fueled by two articles that showed that many publications related to a new cyclooxygenase (COX) inhibitor attributed primarily or solely to academic investigators were actually written by the sponsoring pharmaceutical company or medical publishing companies hired by the company (Ross et al. 2008) and that the sponsoring pharmaceutical company manipulated the data analysis in two clinical trials to minimize the increased mortality associated with the drug (Psaty and Kronmal 2008). According to several authors, bias in the way industry-sponsored research is conducted and reported is not unusual (Angell 2004, 2008; DeAngelis and Fontanarosa 2008). Because of the importance of the role of industry-sponsored research in establishing the database often used to generate evidence-based guidelines, we review this issue in detail in this chapter.

Safer (2002) identified several ways that pharmaceutical companies can bias the interpretation of the effects of their medication (see Table 5–7).

TABLE 5–7. Study design issues that can bias the outcome of a study comparing two medications

Using a dose of the comparable drug that is outside of the standard clinical range
Altering the usual dosing schedule of the competing drug
Using misleading research measurement scales
Selecting end points post hoc
Masking unfavorable side effects
Repeatedly publishing the same or similar findings
Selectively highlighting findings favorable to the sponsor
Editorializing in the abstract
Publishing the obvious
Engaging in statistical obfuscation
Selecting subjects and a time frame designed to achieve a favorable outcome
Withholding unfavorable results
Using masked sponsorship

Source. Adapted from Safer 2002.

Using Doses Outside the Usual Range for Competitive Advantage

If a new medication is compared with an older medication prescribed at a higher-than-usual dose and a dose associated with more side effects, then patients taking the older medication will be less likely to continue taking it. For instance, Safer (2002) noted that at least eight industry-sponsored trials have compared a second-generation neuroleptic drug to a fixed high dose of haloperidol (20 mg) or to a haloperidol dose averaging >20 mg/day. Doses of haloperidol exceeding the customary levels of 4–10 mg/day may not produce better clinical results than do doses greater than that range, but they induce more extrapyramidal side effects (EPS) and lead to far more treatment dropouts. Comparing a newer neuroleptic with an older one used at a very high dose also can make the newer medication appear to have lower rates of EPS. However, one might argue that a high dose of the medication used for comparison may ensure that patients receive an adequate dose.

Substantially Altering the Dose Schedule of the Comparison Drug for Competitive Advantage

Industry-sponsored studies comparing two antidepressant drugs may schedule an unusually rapid and substantial dose increase in the one not manufactured by the sponsoring company (Safer 2002). Doses and dose schedules beyond the usual range, particularly early in treatment, may bring an increased rate of side effects. In studies comparing side effects of a new medication with those of an older one, use of a relatively low dose of the new medication may minimize the side effects of the new medication compared with the older one. In judging any industry-sponsored trial, the reader should assess whether the protocol follows standard dosing and if not, why not.

Using Self-Serving Measurement Scales and Making Misleading Conclusions From Measurement Findings

In early studies of risperidone, the investigations used a measure of the "worst extrapyramidal symptoms score," which Safer (2002) argued creates the impression that the medication was not associated with EPS. Alternatively, a nonstandard criterion for interpreting the outcome from a measurement can be misleading.

Selecting the Major Findings and End Points Post Hoc

Safer (2002) noted that some industry-sponsored trials select the end points post hoc. Doing so opens the door to "fishing" for results favorable to the medication. For instance, Jureidini and Tonkin (2003) argued that in a study of the effects of paroxetine on reducing depression, the authors changed the predetermined criteria to a criterion that seemed to favor the medication. However, Keller et al. (2003) responded to Jureidini and Tonkin's criticism by arguing that the field had since considered some of the secondary measures as better than those initially proposed. The CONSORT guidelines require that the primary analyses be specified in advance and followed up in part to avoid such problems. Nevertheless, secondary, exploratory analyses can provide important clues that need to be confirmed with randomized studies.

Masking Unfavorable Side Effects

Safer (2002) reported that sexual side effects from selective serotonin reuptake inhibitor (SSRI) antidepressants ranged from 2% to 73%, depending primarily on whether side effects were elicited merely by open-ended questioning or by a detailed inquiry. Reports by sponsoring pharmaceutical companies may downplay sexual side effects of SSRIs by using open-ended or nonspecific questions about side effects (Zajecka et al. 1999). In one instance, a drug company–sponsored review including more than 3,000 patients receiving SSRIs simply did not list any sexual side effects on its 23-item side-effect table (Preskorn 1997).

Publishing the Same or Similar Results

Safer (2002) noted that pharmaceutical manufacturers may publish the same or similar positive study results, perhaps to increase the visibility of their product. Such practices can be misleading and even have a detrimental effect on one of the key evidence-based medicine practices, the use of meta-analyses. For instance, Huston and Moher (1996) performed a meta-analysis of the effects of the antipsychotic agent risperidone. They identified 20 articles and several unpublished reports describing randomized, double-blind trials, but after a search they described as "vexing," "bewildering," and "intolerably time-consuming," they concluded that there were probably only 7 small trials and 2 large trials, one of the latter being reported "in part, transparently, and not so transparently in six different publications…the authorship was different for each." They wrote: "Multiple renditions of the same information is self-serving, wasteful, abuses the volunteer time of peer reviewers, and can be profoundly misleading; it brings into question the integrity of medical research." An editorial in *JAMA* by Rennie (1999) discussed in depth this problem of publication duplication.

Other Biases

Other biases in industry-sponsored trials include, as noted by Safer (2002), editorializing in the abstract, publishing the obvious, not following standard statistical procedures, including "borderline" significant findings, designing studies that select subjects and a time frame to achieve a favorable outcome, and withholding unfavorable results. The role of masked sponsorship was noted earlier.

Systematic Reviews of Industry-Sponsored Studies

Several reviews of industry-sponsored trials have suggested that there is a systematic bias favoring products that are made by the company funding the research. In one recent review, the authors did not find that studies funded by industry were of poorer quality than those funded by other sources of support, but studies sponsored by pharmaceutical companies were more likely to have outcomes favoring the company's product than were studies with other sponsors (OR=4.05, 95% CI 2.98–5.51; 18 comparisons) (Lexchin et al. 2003). Montgomery et al. (2004) examined randomized clinical trials of second-generation antipsychotics for treating schizophrenia. They noted that "within the industry-funded studies, outcomes of trials involving first authors employed by industry sponsors demonstrated a trend toward second generation over first generation antipsychotic to a greater degree than did trials involving first authors employed outside the industry (P=0.05)" (Montgomery et al. 2004).

Reporting study bias is not limited, of course, to the pharmaceutical industry. In the above-mentioned study by Montgomery et al. (2004) questioning industry-sponsored trials, the authors committed several of the errors noted by Safer (2002). They say, for instance, "Non-industry-funded studies showed a trend toward higher quality than industry-funded studies; however, the difference between the two was not significant."

In another analysis of the effects of funding source on outcome, Heres et al. (2006) identified 42 reports of head-to-head comparisons of second-generation antipsychotics. Of these, 33 were sponsored by a pharmaceutical company. In 90% of the studies, the reported overall outcome was in favor of the sponsor's drug. The authors noted that "This pattern resulted in contradictory conclusions across studies when the findings of studies of the same drugs but with different sponsors were compared." Heres and colleagues recommended that peer reviewers should verify whether the abstract really summarizes the overall results of the trial. They noted that in a study sponsored by the manufacturer of olanzapine that compared olanzapine with risperidone, the two medications were not different on 21 of 25 efficacy measures, yet the abstract emphasized the greater efficacy of olanzapine. In contrast, in a study sponsored by the manufacturer of risperi-

done that compared the same set of agents (risperidone and olanzapine), the two medications were found to be not different on 33 of 37 efficacy measures, including the a priori primary end points of the study, yet the abstract emphasized the greater efficacy of risperidone (Tandon 2006).

Caveat Emptor

In recent years, the rise of clinical trial registries, the requirement that studies meet prespecified standards to be included in medical reviews, the adoption of the CONSORT standards, and the emphasis by medical publishers to address issues like those raised earlier may have reduced potential bias in trials, including those sponsored by the pharmaceutical industry. Nevertheless, the practitioner should be very alert to the potential for bias, particularly given the influence of pharmaceutical representatives in medical education. Dr. Daniel Carlat (2007) described how he had been recruited, as a respected practitioner, to give drug talks and how difficult he found it to question the evidence he had been given to present.

Antidepressants and Suicide Risk

The recent controversy about the potential risk of antidepressants increasing suicidal ideation or risk in adolescents illustrates some of the complex issues related to evaluating evidence. We include it here, in part because the role of the pharmaceutical industry has been highlighted, but also to illustrate many of the limitations of the knowledge base used in evidence-based psychiatry practice.

The possibility that antidepressant medications, especially SSRIs, increase the risk of suicidal behavior was first raised in several case reports of children and adults during the early 1990s (King et al. 1991; Rothschild and Locke 1991; Teicher et al. 1990). In 2003, the U.K. Department of Health warned physicians against prescribing any SSRI antidepressant drug except fluoxetine for depressed youths younger than 18 years. In 2003, the FDA was also concerned about this risk and announced that it was reviewing "a possible increased rate" of suicidal behavior in youths taking paroxetine hydrochloride (U.S. Food and Drug Administration 2003). The FDA recommended that paroxetine not be used in children and adolescents for the treatment of major depressive disorder. On October 15, 2004, the FDA issued a so-called black-box warning that all antidepressants pose significant risks of suicidality in children and adolescents and that children and adults taking antidepressants should be watched closely for increased suicidal thinking or behavior (U.S. Food and Drug Administration 2004). The warning immediately generated controversy in addition to opening discussion on more general issues of the FDA's role in drug safety.

One researcher believed that the focus for reporting should be on the pharmaceutical industry:

> Why was it left to regulatory bodies to publicize the lack of effectiveness of paroxetine and venlafaxine? The single published placebo controlled trial concluded that paroxetine was effective and safe in adolescent depression. But none of the large negative trials (2 each for paroxetine and venlafaxine) were published, a phenomenon that undermines evidence-based medicine. Pharmaceutical companies seeking regulatory approval are obliged to make the results of all clinical trials they sponsor available to regulatory agencies. However, there is no requirement for these results to be published or even made available to investigators. Those researchers, including myself, who *did* see results of negative paroxetine industry trials were prohibited by nondisclosure contracts from discussing them. (Garland 2004)

The black-box warning was controversial from the outset. Several organizations and researchers began to re-review both the efficacy and the risk of antidepressants used in children and adolescents. In 2003, the American College of Neuropharmacology (ACNP) convened a task force to evaluate the evidence for safety of SSRIs in youths, including reviewing published clinical trials and data from the U.K. Medicines and Healthcare Products Regulatory Agency (MHRA), as well as FDA analyses and reports made public online (Hammad 2004). The review by the ACNP concluded that "fluoxetine is effective in treating depression in children and adolescents" (Mann et al. 2006). The task force also addressed the relation between SSRIs and suicidal behavior in youths. The FDA's meta-analyses of the adverse event report data pooled across all drugs indicated that there was a statistically significant elevation (about twofold) in risk of suicidality for antidepressants relative to placebo (Hammad 2004). However, the ACNP report did not find convincing evidence that SSRIs increased the suicide rate in

youths. The World Psychiatric Association (Möller et al. 2008) undertook an extensive review of the database and concluded that "antidepressants, including SSRIs, carry a small risk of inducing suicidal thoughts and suicide attempts, in age groups below 25 years."

Klein (2006) argued that the FDA relied on a poorly defined "surrogate measure" of completed suicide and noted that the validation of surrogate measures is very difficult when used to predict very rare events. He also noted that a major potential source of information of medications—that is, of their use "postmarketing"—is deeply flawed because it can rely on a few events (such as completed suicides) without considering the more complicated epidemiologic issues that would provide a more accurate position of risk (e.g., what is the relative risk in the population).

Posner et al. (2007) have developed the Columbia Classification Algorithm of Suicide Assessment (C-CASA) to classify "suicidal events" in terms of suicidal ideation, preparation toward suicide attempt, or completed suicide. In an analysis of 25 pediatric antidepressant trials, they found that the system was reliable. The FDA safety analysis of the risk of suicidality in a depressed pediatric sample that used these C-CASA ratings (Hammad et al. 2006) found reduced risk estimates when compared with earlier FDA estimates that relied on the pharmaceutical label. The practitioner faced with the "evidence" encounters a very complicated and confusing issue.

The warnings had an effect on reducing antidepressant use (Olfson et al. 2008). During the prewarning study period, a 36.0% per year ($P<0.001$) increase in total youth (ages 6–17 years) antidepressant use was reported, which was followed by decreases of 0.8% per year ($P=0.85$) and 9.6% per year ($P=0.21$) during the paroxetine and black-box warning study periods, respectively. The difference in trends between the prewarning and the paroxetine warning periods was significant ($P<0.001$). Paroxetine use in young people also significantly increased during the prewarning study period (30.0% per year; $P<0.001$) before significantly declining during the paroxetine warning study period (–44.2% per year; $P<0.001$) (Olfson et al. 2008).

There has been a more recent concern about raised suicide rates in adolescents (Bridge et al. 2008). For males and females, ages 10–19, rates of suicide increased from 2003 to 2004. Although for males, the suicide rate decreased by 1.8% between 2004 and 2005 (7.13 to 7.00 per 100,000), rates of suicide in 2004 and 2005 were still significantly greater than predicted by the 1996–2003 trend (2004 95% prediction interval [PI] 5.90–6.90; 2005 95% PI 5.63–6.66). The rate of suicide for females ages 10–19 years decreased by 16.7% between 2004 and 2005 (2.22 to 1.85 per 100,000); both 2004 and 2005 rates were significantly greater than the expected rates (2004 95% PI 1.18–1.67; 2005 95% PI 1.11–1.62). Although a reduction in antidepressant use needs to be considered as one factor, Bridge et al. (2008) note:

> Studies to identify causal agents are important next steps. These studies should involve comprehensive assessment of individual-level exposure and outcome data, as aggregate data alone cannot establish causal links. Possible factors to consider include changes in the prevalence of known risk factors (e.g., alcohol use, access to firearms), the influence of Internet social networks, higher rates of untreated depression in the wake of recent boxed warnings on antidepressants and increases in suicide among US troops, some being older adolescents.

The ACNP task force recommended the continued use of fluoxetine as an effective and readily available treatment for major depression in youth suicide, suicidal thinking, and plans for suicide. They also recommended that ongoing monitoring of suicidal thoughts in patients taking antidepressants is necessary.

The American Academy of Child and Adolescent Psychiatry (AACAP) urged the FDA not to issue a black-box warning against the use of all antidepressants for the treatment of depression in children and adolescents, noting that "Efficacy and safety data on pediatric antidepressant use has been the subject of ongoing review. The research and its reviews show efficacy, while the signal for the risks of increased suicidal thinking and self-harm events is not strong and can be monitored" (Sarles 2004). AACAP and the American Psychiatric Association (n.d.) also issued guidelines for physicians related to this recommendation.

What is the truth about the risk-benefit ratio of antidepressant use in depressed children and adolescents? Are antidepressants beyond fluoxetine effective in reducing depressive symptoms in children? In adolescents? Do they increase suicidal thoughts?

Behaviors? Completed events? Is the reduction in antidepressant use responsible for the rise in suicide rates in recent years? The uncertainty of the answers to these questions reflects the limits of evidence-based practice and the challenges to the practitioner in using evidence-based recommendations.

References

Altman DG: Practical Statistics for Medical Research. Boca Raton, FL, Chapman & Hall, 1991

Altman DG: Confidence intervals for the number needed to treat. BMJ 317:1309–1312, 1998

Altman DG, Doré CJ: Randomisation and baseline comparisons in clinical trials. Lancet 335:149–153, 1990

Altman DG, Schulz KF: Statistics notes: concealing treatment allocation in randomised trials. BMJ 323:446–447, 2001

Altman DG, Schulz KF, Moher D, et al: The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 134:663–694, 2001

American Psychiatric Association (APA), American Academy of Child and Adolescent Psychiatry (AACAP): The use of medication in treating childhood and adolescent depression: information for physicians. n.d. Available at: http://www.parentsmedguide.org/physiciansmedguide.pdf. Accessed June 8, 2009

Angell A: The Truth About the Drug Companies: How They Deceive Us and What to Do About It. New York, Random House, 2004

Angell A: Industry-sponsored clinical research: a broken system. JAMA 300:1069–1071, 2008

Badenoch D, Heneghan C: Evidence-Based Medicine Toolkit. London, BMJ Books, 2002

Begg C, Cho M, Eastwood S, et al: Improving the quality of reporting of randomized controlled trials: the CONSORT statement. JAMA 276:637–639, 1996

Bland M: An Introduction to Medical Statistics, 3rd Edition. Oxford, UK, Oxford University Press, 2000

Bridge JA, Greenhouse JB, Weldon AH, et al: Suicide trends among youths aged 10 to 19 years in the United States, 1996–2005. JAMA 300:1025–1026, 2008

Buscemi N, Vandermeer B, Friesen C, et al: The efficacy and safety of drug treatments for chronic insomnia in adults: a meta-analysis of RCTs. J Gen Intern Med 22:1335–1350, 2007

Carlat D: Dr. Drug Rep. The New York Times Magazine, November 25, 2007

Chaput de Saintonge DM, Herxheimer A: Harnessing placebo effects in health care. Lancet 344:995–998, 1994

Charney DS, Nemeroff CB, Lewis L, et al: National Depressive and Manic-Depressive Association consensus statement on the use of placebo in clinical trials of mood disorders. Arch Gen Psychiatry 59:262–270, 2002

Citrome L: Compelling or irrelevant? Using number needed to treat can help decide. Acta Psychiatr Scand 117:412–419, 2008

Collins R, MacMahon S: Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. Lancet 357:373–380, 2001

Cook RJ, Sackett DL: The number needed to treat: a clinically useful measure of treatment effect. BMJ 310:452–454, 1995

Cookson J, Taylor D, Katona C: The Use of Drugs in Psychiatry: The Evidence From Psychopharmacology, 5th Edition. London, Gaskell, 2002

Crow R, Gage H, Hampson S, et al: The placebo effect: methodological process and implications of a structured review, in The Advanced Handbook of Methods in Evidence Based Healthcare. Edited by Stevens A, Abrams K, Brazier J, et al. London, Sage, 2001, pp 73–89

Cummings SR, Grady D, Hulley SB: Designing an experiment: clinical trials I, in Designing Clinical Research, 2nd Edition. Edited by Hulley SB, Cummings SR, Browner WS, et al. Philadelphia, PA, Lippincott Williams & Wilkins, 2001, pp 143–155

Daly LE, Bourke GJ: Interpretation and Uses of Medical Statistics, 5th Edition. Oxford, UK, Blackwell Scientific, 2000

DeAngelis CD, Fontanarosa PB: Impugning the integrity of medical science: the adverse effects of industry influence. JAMA 299:1833–1835, 2008

Deeks JJ, Altman DG: Effect measures for meta-analysis of trials with binary outcomes, in Systematic Reviews in Health Care: Meta-Analyses in Context, 2nd Edition. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 313–335

Devereaux PJ, Manns BJ, Ghali WA, et al: Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. JAMA 285:2000–2003, 2001

Devereaux PJ, Bhandari M, Montori VM, et al: Double blind, you have been voted off the island! Evid Based Ment Health 5:36–37, 2002

Easterbrook PJ, Berlin JA, Gopalan R, et al: Publication bias in clinical research. Lancet 337:867–872, 1991

Everitt BS: Statistical Methods for Medical Investigations. New York, Oxford University Press, 1989

Fletcher RH, Fletcher SW, Wagner EH: Clinical Epidemiology: The Essentials, 3rd Edition. Baltimore, MD, Williams & Wilkins, 1996

Furukawa TA, Guyatt GA, Griffith LE: Can we individualize the "number needed to treat"? An empirical study of summary effect measures in meta-analyses. Int J Epidemiol 31:72–76, 2002

Gardner MJ, Altman DG: Confidence intervals rather than P values, in Statistics With Confidence, 2nd Edition. Edited by Altman DG, Machin D, Bryant TN, et al. London, BMJ Books, 2000, pp 15–27

Garland EJ: Facing the evidence: antidepressant treatment in children and adolescents. CMAJ 170:489–491, 2004

Geddes J, Butler R: Depressive disorders. Clin Evid 6:726–742, 2001

Gray GE: Evidence-based medicine: an introduction for psychiatrists. J Psychiatr Pract 8:5–13, 2002

Gray GE: Concise Guide to Evidence-Based Psychiatry. Washington, DC, American Psychiatric Publishing, 2004

Greenhalgh T: How to Read a Paper: The Basics of Evidence Based Medicine, 2nd Edition. London, BMJ Books, 2001

Guyatt G: Therapy and harm: why study results mislead—bias and random error, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 223–231

Guyatt G, Devereaux PJ: Therapy and validity: the principle of intention-to-treat, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 267–273

Guyatt G, Cook D, Devereaux PJ, et al: Therapy, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002a, pp 55–79

Guyatt G, Haynes B, Jaeschke R, et al: Introduction: the philosophy of evidence-based medicine, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002b, pp 3–12

Guyatt G, Walter S, Cook D, et al: Therapy and understanding the results: confidence intervals, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002c, pp 339–349

Halpern SD, Karlawish JHT, Berlin JA: The continuing unethical conduct of underpowered clinical trials. JAMA 288:358–362, 2002

Hammad TA: Review and evaluation of clinical data. August 16, 2004. Available at: http://www.fda.gov/ohrms/dockets/ac/04/briefing/2004-4065b1-10-TAB08-Hammads-Review.pdf. Accessed June 9, 2009

Hammad TA, Laughren T, Racoosin J: Suicidality in pediatric patients treated with antidepressant drugs. Arch Gen Psychiatry 63:332–333, 2006

Hay P, Bacaltchuk J: Bulimia nervosa. Clin Evid 6:715–725, 2001

Hennekens CH, Buring JE, Mayrent SL: Epidemiology in Medicine. Boston, MA, Little, Brown, 1987

Heres S, Davis J, Maino K, et al: Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: an exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. Am J Psychiatry 163:185–194, 2006

Holden JD: Hawthorne effects and research into professional practice. J Eval Clin Pract 7:65–70, 2001

Huston P, Moher D: Redundancy, disaggregation, and the integrity of medical research. Lancet 347:1024–1026, 1996

Ioannidis JPA, Haidich A-B, Pappa M, et al: Comparison of evidence of treatment effects in randomized and nonrandomized studies. JAMA 286:821–830, 2001

Jadad A: Randomised Controlled Trials. London, BMJ Books, 1998

Jaeschke R, Guyatt G, Barratt A: Therapy and understanding the results: measures of association, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 351–368

Jüni P, Altman DG, Egger M: Systematic reviews in health care: assessing the quality of controlled clinical trials. BMJ 323:42–46, 2001

Jureidini J, Tonkin A: Paroxetine in major depression (letter). J Am Acad Child Adolesc Psychiatry 42:514, 2003

Kaptchuk TJ: Powerful placebo: the dark side of the randomised clinical trial. Lancet 351:1722–1725, 1998

Keller MB, Ryan R, Strober M, et al: Paroxetine in major depression: Dr. Keller et al reply. J Am Acad Child Adolesc Psychiatry 42:514–515, 2003

King RA, Riddle MA, Chappell PB, et al: Emergence of self-destructive phenomena in children and adolescents during fluoxetine treatment. J Am Acad Child Adolesc Psychiatry 30:179–186, 1991

Klein DF: The flawed basis for FDA post-marketing safety decisions: the example of anti-depressants and children. Neuropsychopharmacology 31:689–699, 2006

Kristiansen IS, Gyrd-Hansen D, Nexoe J, et al: Number needed to treat: easily understood and intuitively meaningful? Theoretical considerations and a randomized trial. J Clin Epidemiol 55:888–892, 2002

Kunz R, Oxman AD: The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. BMJ 317:1185–1190, 1998

Lacchetti C, Guyatt G: Therapy and validity: surprising results of randomized controlled trials, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 247–265

Laporte J-R, Figueras A: Placebo effects in psychiatry. Lancet 344:1206–1209, 1994

Laupacis A, Sackett DL, Roberts RS: An assessment of clinically useful measures of the consequences of treatment. N Engl J Med 318:1728–1733, 1988

Lexchin J, Bro LA, Djulbegovic B, et al: Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ 326:1167–1170, 2003

Mann JJ, Emslie G, Baldessarini R, et al: ACNP Task Force report on SSRIs and suicidal behavior in youth. Neuropsychopharmacology 31:473–492, 2006

McIntosh A, Lawrie S: Schizophrenia. Clin Evid 6:776–797, 2001

Moher D, Schulz KF, Altman D: The CONSORT statement: revised recommendation for improving the quality of reports of parallel-group randomized trials. JAMA 285:1987–1991, 2001

Möller HJ, Baldwin DS, Goodwin G, et al: Do SSRIs or antidepressants in general increase suicidality? WPA Section on Pharmacopsychiatry: consensus statement. Eur Arch Psychiatry Clin Neurosci 258 (suppl 3):3–23, 2008

Montgomery JH, Byerly M, Carmody T, et al: An analysis of the effect of funding source in randomized clinical trials of second generation antipsychotics for the treatment of schizophrenia. Control Clin Trials 25:598–612, 2004

Montori V, Guyatt G: Summarizing the evidence: publication bias, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 529–538

Montori VM, Bhandari M, Devereaux PJ, et al: In the dark: the reporting of blinding status in randomized controlled trials. J Clin Epidemiol 55:787–790, 2002

Nuovo J, Melnikow J, Chang D: Reporting number needed to treat and absolute risk reduction in randomized controlled trials. JAMA 287:2813–2814, 2002

Olfson M, Marcus SC, Druss BG: Effects of Food and Drug Administration warnings on antidepressant use in a national sample. Arch Gen Psychiatry 65:94–101, 2008

Perlis RH, Perlis CS, Wu Y, et al: Industry sponsorship and financial conflict of interest in the reporting of clinical trials in psychiatry. Am J Psychiatry 162:1957–1960, 2005

Peto R, Pike MC, Armitage P, et al: Design and analysis of randomized clinical trials requiring prolonged observation of each patient, I: introduction and design. Br J Cancer 34:585–612, 1976

Pocock SJ, Hughes MD, Lee RJ: Statistical problems in the reporting of clinical trials: a survey of three medical journals. N Engl J Med 317:426–432, 1987

Posner K, Oquendo MA, Gould M, et al: Columbia Classification Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants. Am J Psychiatry 164:1035–1043, 2007

Preskorn SH: Clinically relevant pharmacology of selective serotonin reuptake inhibitors: an overview with emphasis on pharmacokinetics and effects on oxidative drug metabolism. Clin Pharmacokinet 32 (suppl 1):1–21, 1997

Psaty BM, Kronmal RA: Reporting mortality findings in trials of rofecoxib for Alzheimer disease or cognitive impairment: a case study based on documents from rofecoxib litigation. JAMA 299:1813–1817, 2008

Rennie D: Fair conduct and fair reporting of clinical trials. JAMA 282:1766–1768, 1999

Rosenthal R, Jacobson L: Pygmalion in the Classroom. New York, Holt, Rinehart & Winston, 1968

Rosenthal R, Rosnow RL: Essentials of Behavioral Research: Methods and Data Analysis, 2nd Edition. Boston, MA, McGraw-Hill, 1991

Ross JS, Hill KP, Egilman DS, et al: Guest authorship and ghostwriting in publications related to rofecoxib: a case study of industry documents from rofecoxib litigation. JAMA 299:1800–1812, 2008

Rothschild AJ, Locke CA: Reexposure to fluoxetine after serious suicide attempts by three patients: the role of akathisia. J Clin Psychiatry 52:491–493, 1991

Sackett DL, Haynes RB, Guyatt GH, et al: Clinical Epidemiology: A Basic Science for Clinical Medicine, 2nd Edition. Boston, MA, Little, Brown, 1991

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Safer DJ: Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. J Nerv Ment Dis 190:583–592, 2002

Sarles RM: AACAP comment on pending pediatric antidepressants warning. September 28, 2004. Available at: http://www.aacap.org/cs/root/resources_for_families/psychiatric_medication/aacap_comment_on_pending_pediatric_antidepressants_warning. Accessed June 8, 2009

Schulz KF: Assessing allocation concealment and blinding in randomised controlled trials: why bother? Evid Based Ment Health 3:4–5, 2000

Sitthi-amorn C, Poshyachinda V: Bias. Lancet 342:286–288, 1993

Song F, Eastwood A, Gilbody S, et al: Publication and related biases, in The Advanced Handbook of Methods in Evidence Based Healthcare. Edited by Stevens A, Abrams K, Brazier J, et al. London, Sage, 2001, pp 371–390

Sterne JAC, Smith GD: Sifting the evidence: what's wrong with significance tests? BMJ 322:226–231, 2001

Straus SE, McAlister F: Applying the results of trials and systematic reviews to our individual patients. Evid Based Ment Health 4:6–7, 2001

Streiner D, Geddes J: Intention to treat analysis in clinical trials when there are missing data. Evid Based Ment Health 4:70–71, 2001

Szatmari P: Some useful concepts and terms used in articles about treatment. Evid Based Ment Health 1:39–40, 1998

Tandon R: Comparing antipsychotic efficacy. Am J Psychiatry 163:1645, 2006

Teicher MH, Gold C, Cole JO: Emergence of intense suicidal preoccupation during fluoxetine treatment. Am J Psychiatry 147:207–210, 1990

U.S. Food and Drug Administration: FDA statement regarding the anti-depressant Paxil for pediatric population (FDA Talk Paper). June 19, 2003. Available at: http://www.fda.gov/bbs/topics/ANSWERS/2003/ANS01230.html. Accessed June 8, 2009

U.S. Food and Drug Administration: Worsening depression and suicidality in patients being treated with antidepressant (FDA Public Health Advisory). March 22, 2004. Available at: http://www.fda.gov/Drugs/DrugSafety/PublicHealthAdvisories/ucm161696.htm. Accessed August 11, 2009

Walsh BT, Seidman SN, Sysko R, et al: Placebo response in studies of major depression: variable, substantial, and growing. JAMA 287:1840–1847, 2002

Woods SW, Stolar M, Sernyak MJ, et al: Consistency of atypical antipsychotic superiority to placebo in recent clinical trials. Biol Psychiatry 49:64–70, 2001

Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature: Call for comments on a proposal to improve reporting of clinical trials in the biomedical literature. Ann Intern Med 121:894–895, 1994

Yudkin PL, Stratton IM: How to deal with regression to the mean in intervention studies. Lancet 347:241–243, 1996

Zajecka J, Amsterdam JD, Quitkin FM, et al: Changes in adverse events reported by patients during 6 months of fluoxetine therapy. J Clin Psychiatry 60:389–394, 1999

This page intentionally left blank

6

Systematic Reviews and Meta-Analyses

Gregory E. Gray, M.D., Ph.D.

In Chapter 5, Taylor and I discussed individual studies of therapies. As we pointed out in that chapter, both false-positive and false-negative results can occur in clinical trials because of chance and bias. In particular, false-negative results are common because studies are often too small to detect important treatment effects. As a result, when attempting to answer a clinical question, it is best not to depend on the results of a single clinical trial. It is preferable instead to "average out" the results of all such clinical trials related to a specific clinical question because this will give a better estimate of the treatment effect than will the results of any single study. This is exactly what is done in a meta-analysis, and it is why systematic reviews rank at the top of the evidence hierarchy (Table 4–1, Chapter 4).

Narrative Reviews Versus Systematic Reviews

In the typical *journalistic* or *narrative* review article, the author attempts to present a coherent review of a topic, selectively citing the literature to support the statement made in the article. When studies of a treatment are not in agreement, the author of the review may tally the positive and negative results (sometimes called *vote counting*), indicating that controversy exists and that more research is needed to resolve the issue.

Although such reviews are common in medical journals (Rochon et al. 2002), they are often misleading, reflecting the author's biases in selectively reviewing the literature (D.J. Cook et al. 1998; Egger et al. 2001c; Greenhalgh 2001). Such reviews may contribute to delays in implementing changes in clinical practice that are based on important research findings (Egger et al. 2001c).

TABLE 6–1. Steps in conducting a systematic review

1. Formulate the question
2. Locate studies
Online databases (e.g., MEDLINE, EMBASE)
Registers of clinical trials:
ClinicalTrials.gov
Cochrane Central Register of Controlled Trials
*meta*Register of Controlled Trials
Contact authors or manufacturers
Check reference lists
Perform manual searches
3. Assess study quality
Rating scales
Two or more reviewers
4. Extract and summarize the data
Tables
Forest plots
Pooled effect size and confidence intervals

The systematic review is a better alternative to the traditional review (D.J. Cook et al. 1998; Egger et al. 2001c; Greenhalgh 2001). Such a review focuses on a specific clinical question, involves a comprehensive literature search, and often combines the study results mathematically through meta-analytic techniques (D.J. Cook et al. 1998).

Conducting a Systematic Review

The steps involved in conducting a systematic review are summarized in Table 6–1 and are described briefly in the following subsections. Readers who want more detail about the conduct of systematic reviews are referred to several textbooks (Clarke and Oxman 2002; Egger et al. 2001b; Glasziou et al. 2001). The National Health Service Centre for Reviews and Dissemination (2001) publishes a guide for systematic reviews that can be found at www.york.ac.uk/inst/crd/report4.htm.

Formulate the Question

The first step in the process is formulating the question. The process is similar to that of the evidence-based medicine (EBM) model (see Chapter 2) and involves a similar 4-part PICO question (patient/problem, intervention, comparison, and outcome; see Chapter 3).

Locate Studies

The second step in the process involves finding all of the relevant studies. Such a search involves using the online databases described in Chapter 4, but using filters for *sensitivity* (rather than for *specificity*) (Glasziou et al. 2001; Robinson and Dickersin 2002). Such a strategy will, however, miss many early (pre-1990) studies, non-English-language studies, and unpublished studies, so additional effort may be needed to identify such studies (Egger and Smith 1998; Egger et al. 2001a; Glasziou et al. 2001; Lefebvre and Clarke 2001). Methods that authors undertake to identify such studies include contacting authors of published studies or manufacturers of drugs, checking reference lists of published studies or prior reviews ("snowballing"), manually searching journals or proceedings abstracts that are not abstracted in MEDLINE or similar databases, and checking databases of clinical trials (Glasziou et al. 2001; Helmer et al. 2001; Lefebvre and Clarke 2001). The latter include the Cochrane Central Register of Controlled Trials (available at www.mrw.interscience.wiley.com; also as part of Ovid's Evidence-Based Medicine Reviews database) and the *meta*Register of Controlled Trials (www.controlled-trials.com/mrct). If such measures are not taken, several sources of bias may influence the results of the review, including publication bias and language bias (Table 6–2).

Some meta-analyses rely on unpublished data or data requested from an investigator. Because negative studies may be more difficult to publish than positive studies (Dwan et al. 2008), a careful acquisition of unpublished data from well-conducted studies would allow for a more accurate appraisal of the effects of an intervention than simply using published studies. For instance, Turner et al. (2008) obtained reviews from the U.S. Food and Drug Administration (FDA) for studies of 12 antidepressant agents involving 12,564 patients. They then tried to identify matching publications in the medical literature. The effect size derived from the published reports was compared with the effect size derived from the entire FDA data set. About 74 FDA-registered studies (31%), accounting for 3,449 study participants, were not published. A total of 37 studies reviewed by the FDA as having positive results were published; only 1 study viewed as positive was not published. Studies viewed by the FDA as having negative or questionable results were, with three exceptions, either not published (22 studies) or published in a way that, in the opinion of Turner et al. (2008), conveyed a positive outcome. According to the published literature, it appeared that 94% of the trials conducted were positive. By contrast, the FDA analysis showed that 51% were positive. Separate meta-analyses of the FDA and journal data sets showed that the increase in effect size ranged from 11% to 69% for individual drugs and was 32% overall. One cannot assume that a review of published pharmaceutical studies represents the best estimate of the effect of the medication.

Publication bias, also known as the *file drawer problem*, refers to a tendency to preferentially publish "positive" results (Egger and Smith 1998; Egger et al. 2001a; Montori and Guyatt 2002; Montori et al. 2000; Rosenthal 1979; Sterne et al. 2001). Although it is clear that publication bias occurs, the reasons for it are a topic of some debate, but they include decisions by the investigators, journal editors and reviewers, and pharmaceutical company influences (Egger and Smith 1998; Egger et al. 2001a; Montori

TABLE 6–2. Types of reporting bias in systematic reviews

Type of bias	Description
Publication bias	Results are more apt to be published if they are significant.
Time lag bias	Significant results are published sooner than nonsignificant results.
Language bias	Significant results are submitted to English-language journals, nonsignificant results to non-English-language journals.
Database bias	Studies with significant results are more likely to be published in a journal that is indexed in a database.
Citation bias	Likelihood of article being cited depends on results.
Duplicate publication bias	Results of study appear in more than one publication.
Outcome reporting bias	Selective reporting of some study results.

Source. Adapted from Egger and Smith 1998; Egger et al. 2001a.

and Guyatt 2002; Montori et al. 2000; Olson et al. 2002; Song et al. 2001; Stern and Simes 1997; Thornton and Lee 2000).

Language bias can occur if English-language journals publish a greater proportion of positive studies than do non-English-language journals. Such a bias could occur if foreign investigators preferentially submit positive findings to English-language journals (Egger and Smith 1998; Egger et al. 2001a; Song et al. 2001). *Location bias*, a similar type of bias, refers to the tendency to publish lower-quality positive studies of complementary medicine in low-impact journals (Pittler et al. 2000). The extent to which language and location bias occurs seems to vary, depending on the medical specialty and the disease in question (Jüni et al. 2002; Moher et al. 2000b).

Assess Study Quality

Because the goal of the search strategy is to be as inclusive as possible to avoid missing relevant studies, a search typically will identify numerous studies, many of which are either irrelevant or of poor quality. Including such irrelevant or poor-quality studies in the review and subsequent meta-analysis yields misleading results. Statisticians refer to this as "garbage in, garbage out."

As was discussed in Chapter 5, evidence from less rigorously designed studies is more apt to be biased and misleading than is evidence from more rigorously designed studies (Greenhalgh 2001; Guyatt 2002; Lacchetti and Guyatt 2002; Schulz et al. 1995). It is therefore necessary to appraise the studies identified in the search and to limit further analysis to those studies that meet certain quality criteria (Jüni et al. 2001). The method of doing so is similar to the approach used in Chapter 5 to appraise a randomized controlled trial (RCT). A standardized approach to appraisal, generally involving two reviewers, should be used to avoid biasing the review (Glasziou et al. 2001; Moher et al. 2001). A variety of instruments for rating the quality of studies have been developed for this purpose, but there is no single best instrument (Jüni et al. 2001; West et al. 2002).

Specific study quality aspects may have a greater effect on estimates of treatment effect than others, but this depends on the particular disease and treatment in question (Balk et al. 2002). Nonrandomized trials, trials with inadequate allocation concealment, and unblinded trials tend to overestimate the magnitude of a treatment's effectiveness, although this is not uniformly true for all diseases and their treatments (Altman and Schulz 2001; Ioannidis et al. 2001; Jüni et al. 2001; Kjaergard et al. 2001; Kunz and Oxman 1998; Lacchetti and Guyatt 2002; Schulz 2000). As a general rule, though, selecting higher-quality studies will lead to a less biased and generally less optimistic view of a treatment's effectiveness.

In the process of appraising study quality, an attempt also should be made to identify duplicate or overlapping publications. These can occur, for example, if preliminary data are first published, followed by a second (more complete) publication. It can also occur if different outcome measures are each presented in a separate publication. If such duplicate or overlapping publications are not identified, the same study may be overrepresented in any

meta-analysis, a result known as *multiple publication bias* (Egger and Smith 1998; Tramer et al. 1997).

Extract and Summarize the Data

Once studies meeting the quality criteria are identified, data are abstracted. Such data include methodological details about the study population, intervention, and outcome measures, as well as study results. Such data are typically presented in tabular form in the review, with numerical results often displayed graphically and combined in a meta-analysis. A list of studies that were identified, appraised, and excluded from the review is generally included in the review, along with the reasons for exclusion.

Time and Effort Involved in Systematic Reviews

The process described earlier is conceptually simple; however, it is obviously quite time-consuming, especially if the literature search identifies many studies that then must be assessed for quality. Allen and Olkin (1999) found that systematic review requires 200–2,500 person-hours of effort (median = 1,110 person-hours). About one-half of this effort is spent on the search and retrieval process.

A systematic review is generally a group effort because of the considerable effort involved in its production. Two international collaborative efforts that are involved in producing systematic reviews are the Cochrane Collaboration and the Campbell Collaboration (Antes and Oxman 2001; Cochrane Collaboration 1997; Davies and Boruch 2001).

Meta-Analysis

Meta-analysis refers to the statistical integration of the results of several independent studies (Egger and Smith 1997). In this section, I provide a short nonmathematical overview of meta-analysis. Readers who want an in-depth discussion are referred to several excellent texts (Egger et al. 2001b; Petitti 2000; Sutton et al. 2000).

Meta-analysis involves more than "vote tallying" or merely taking a simple average of the results of various studies. Instead, meta-analysis gives more weight to large studies than to small studies because the results of small studies are subject to more random variability (see Chapter 5). The specific methods by which this is accomplished depend on whether the outcome variable is dichotomous or continuous.

Dichotomous Outcome Measures

As was discussed in Chapter 5, dichotomous outcome measures, such as dying, being readmitted to the hospital, achieving full remission, being rated as "improved" or "much improved," or having at least a 50% decrease in score on a rating scale, are often among the most clinically useful and readily understandable measures. Dichotomous measures are also favored in systematic reviews because of the ease of combining the data in meta-analyses.

Odds Ratio

The measure of treatment effect that is most commonly used in systematic reviews and meta-analyses is the odds ratio (OR). Calculation of the OR was described in Chapter 5. By convention, the outcome measures used in meta-analysis are adverse outcomes; therefore, effective treatments have ORs of 1.0 or less.

The reason that the OR is so frequently used in meta-analyses is that there are several straightforward methods of combining ORs from multiple studies (Deeks et al. 2001; Petitti 2000; Sutton et al. 2000). The two most common methods are the Mantel-Haenszel method (Mantel and Haenszel 1959) and the Peto method (Yusuf et al. 1985). Both methods involve producing a weighted average OR, with larger trials (which have narrower confidence intervals [CIs] for their ORs) being given greater weight than smaller trials (with wider CIs for their ORs), although they differ in the exact weighting used. Both methods assume a "fixed effects" model (see subsection "Fixed and Random Effects Models" later in this chapter). The DerSimonian and Laird (1986) method, used under the "random effects" model, gives somewhat more weight to smaller studies and produces a wider CI for the pooled OR estimate than do the other two methods.

The drawback of using the OR as a measure of treatment effectiveness is that it is not as easily interpretable as relative risk (RR) or number needed to treat (NNT) (Deeks and Altman 2001). Although the OR approximates RR when the frequency of rare outcomes is low (<10%), the two measures diverge as the outcome frequency increases, with the OR overestimating RR (Egger et al. 1997a). Under such cir-

cumstances, misinterpretation of an OR as RR will lead the reader to overestimate or underestimate the effects of a treatment (Deeks and Altman 2001).

Number Needed to Treat

Although NNT is believed by many to be the best expression of relative treatment effectiveness (R.J. Cook and Sackett 1995; Sackett et al. 1991, 2000), its mathematical properties are such that it cannot be directly used in a meta-analysis (Deeks and Altman 2001). NNT can, however, be calculated from the summary OR and the estimated control event rate (CER) with the formula in Appendix B. This approach assumes that the OR for a particular treatment comparison is independent of the CER. Empirically, this appears to be the case (Furukawa et al. 2002; McAlister 2002).

Continuous Outcome Measures

Continuous outcome measures, which include things such as body weight, intelligence quotient, or scores on a rating scale, are common in psychiatric research. In meta-analyses of continuous outcome measures, standardized mean differences, not the means of the outcome measures, are used as the measure of treatment effect (Deeks et al. 2001; Petitti 2000; Sutton et al. 2000). The results of each study are transformed into a standardized mean difference (represented by the letter *d*), using the following formula:

$$d = (\text{mean}_e - \text{mean}_c)/\text{SD}_p$$

where mean_e and mean_c are the means for the experimental and control groups, respectively, and SD_p is the pooled estimate of the standard deviation (SD) of the outcome measure. The results are then pooled by taking a weighted average of the *d*'s for each study, with the weight for a given study being the inverse of the variance of that study's effect size (Deeks et al. 2001; Petitti 2000; Sutton et al. 2000).

The standardized mean difference is a measure of the degree of overlap of the experimental and control group results, expressed in terms of SDs (Freemantle and Geddes 1998). If $d=0$, the means of the control and experimental groups are identical. If $d=1$, the mean of the experimental group is 1 SD above the mean of the control group. From a standard statistical table for a normal distribution, this is equivalent to saying that 84% of the control group has scores below the mean of the experimental group. A table for interpreting the standardized mean difference is given in Appendix B–5.

Fixed and Random Effects Models

As noted earlier, there are both fixed effects and random effects models for combining the results of clinical trials. The choice of model is a topic of debate among statisticians, and readers are referred elsewhere for a more complete discussion (Egger et al. 2001b; Freemantle and Geddes 1998; Montori et al. 2002; Petitti 2000; Sutton et al. 2000). At the risk of oversimplifying a complex issue, it is probably sufficient to say that fixed effects models assume that there is no heterogeneity in study results, whereas random effects models assume that there is heterogeneity. As a result, the random effects model should be preferred when heterogeneity is present.

If there is no heterogeneity, both models will produce similar pooled estimates of the OR, but the CIs obtained with the random effects models are wider (less precise). As a result, the null hypothesis (i.e., that there is no difference between the two treatments) will be rejected less often if calculations are performed with the random effects model. This has led to the claim that the random effects model is overly conservative when minimal heterogeneity is present. In contrast, when there is heterogeneity, hypothesis testing, using the random effects model, gives the more appropriate results. (See Gray and Taylor, Chapter 5 in this volume, for further discussion of hypothesis testing.)

A second practical consideration is that the random effects model gives more equal weighting to large and small studies than does the fixed effects model, which gives more weight to large studies. If there is heterogeneity between the results of large and small studies as the result of either publication bias or differences in quality that vary by study size, the pooled estimate obtained with the random effects model will be more subject to such biases than the estimate obtained with the fixed effects model.

As a practical note, in most cases the two models give similar results. When the results differ, heterogeneity is generally present, and the causes for the heterogeneity need to be taken into account rather than relying on only the pooled estimate of treatment effect.

Forest Plots

Many systematic reviews present their results graphically, in the form of *forest plots* (Egger and Smith 2001; Egger et al. 1997a; Glasziou et al. 2001; Lewis and Clarke 2001; Sutton et al. 2000), also referred to as a *blobbograms* (Freemantle and Geddes 1998). An example of a forest plot is given in Figure 6–1. In this example, the results of each individual study are given by a horizontal bar, with the width of the bar representing the 95% CI for the OR in that study. For each study, a black diamond is used to represent the OR; the size of the diamond is a measure of the size of the study (and hence the weight given to it). Unshaded diamonds are used to indicate the pooled OR, with the width of the diamond indicating the 95% CI for the pooled estimate.

Heterogeneity

Heterogeneity is present when the results from individual studies differ by more than what was expected by chance alone (Freemantle and Geddes 1998; Sutton et al. 2000). Heterogeneity can be assessed either informally (graphically) or through statistical techniques (Sutton et al. 2000).

The informal graphical approach uses forest plots. If there is considerable overlap between the CI bars for the various studies, no heterogeneity is present. Conversely, if the CI bars for some studies do not overlap those of other studies, heterogeneity is present. Figure 6–1 shows that four of the studies found no treatment effect, whereas two studies found the experimental treatment to be significantly better than the control treatment. It can also be seen that the CIs of the four *negative* studies overlap one another and that the CIs of the two *positive* studies also overlap each other. However, the bars for the positive studies do not overlap the Chandler-California study at all; in addition, they barely overlap the Lehman-Baltimore study. Thus, there is heterogeneity in the study results.

The formal statistical approach is to use a chi-square test for heterogeneity. This involves the calculation of a statistic, "Q," which has a chi-square distribution with 1 degree of freedom less than the number of studies. Details of the calculations are given elsewhere (Deeks et al. 2001; Petitti 2000; Sutton et al. 2000).

As noted earlier, when heterogeneity is detected, the random effects model gives a better estimate of pooled effect size than does the fixed effects model. However, the analysis should not stop there. Instead, an attempt should be made to assess the cause of the heterogeneity (Deeks et al. 2001; Glasziou et al. 2001; Petitti 2000; Sutton et al. 2000; Thompson

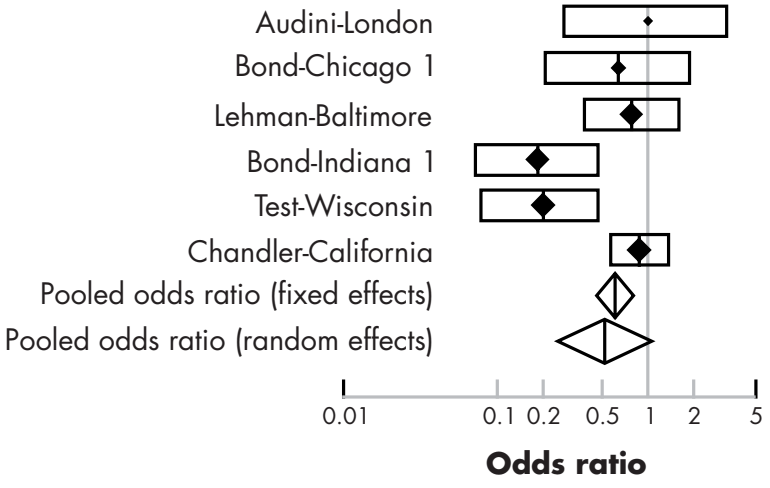


FIGURE 6–1. Example of a forest plot showing effects of assertive community treatment versus usual care on odds of hospitalization of patients with severe mental illness.

Source. Reprinted from Freemantle N, Geddes J: "Understanding and Interpreting Systematic Reviews and Meta-Analyses, II: Meta-Analyses." *Evidence-Based Mental Health* 1:102–104, 1998. Copyright 1998 BMJ Publishing Group. Used with permission.

1994, 2001). Some of the causes of heterogeneity include differences between studies in patient population (e.g., disease severity or other prognostic factors), nature of the intervention (e.g., medication dose, frequency or number of psychotherapy sessions, or presence of co-interventions), compliance, outcome measures, study duration, study quality, and other sources of bias (Glasziou et al. 2001; Sutton et al. 2000).

If the source of the heterogeneity can be identified, it is sometimes useful to present the results by subgroup (Sutton et al. 2000). Unfortunately, differences in subgroups identified in this post hoc fashion may represent the effects of chance or bias rather than true differences in response to treatment (Smith and Egger 2001; Sutton et al. 2000). If, however, such subgroup differences are anticipated from the start of the meta-analysis process (i.e., prior to looking at the forest plot, combining the results, or testing for heterogeneity), then such subgroup analyses may be appropriate.

Funnel Plots and Publication Bias

As previously noted, publication bias is often a concern in systematic reviews. Furthermore, it can bias the results of a meta-analysis, especially if the random effects model is assumed.

Publication bias is usually assessed with a *funnel plot* (Egger et al. 1997b; Glasziou et al. 2001; Montori and Guyatt 2002; Montori et al. 2000; Sterne et al. 2001). In a funnel plot, the *x* axis is treatment effect, and the *y* axis is either study size or a measure of the study standard error, with standard error now considered the preferable measure (Sterne and Egger 2001).

Examples of funnel plots are given in Figure 6–2. Figure 6–2A shows a situation in which no publication bias exists. It can be seen that the plot is symmetrical. On average, the results of smaller studies are the same as those of larger studies; however, results from small studies are more variable. As a result, they deviate more from the mean, producing the funnel shape.

If there is publication bias, there will be a tendency to publish *positive* small study results, but not *negative* small study results. Such a tendency diminishes as the size of the study increases. In Figure 6–2A, the open circles represent the "negative" studies; these studies were not published. Removing these studies gives the asymmetrical funnel plot in Figure 6–2B.

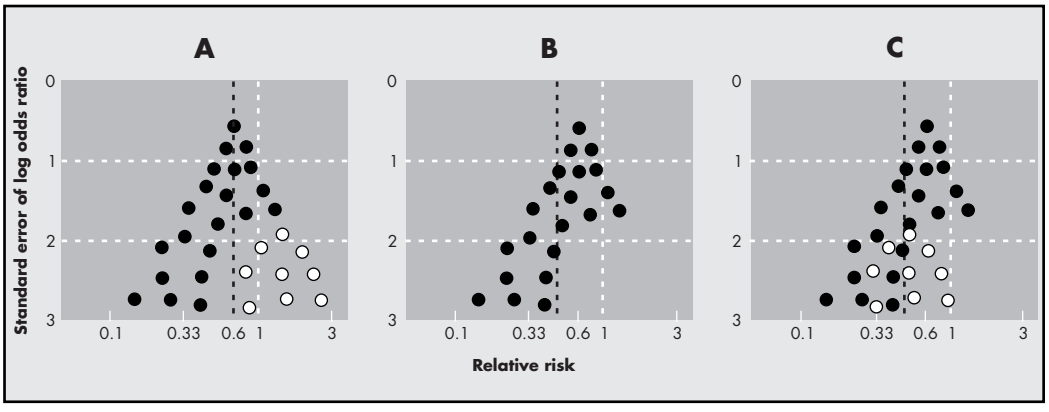


FIGURE 6–2. Examples of funnel plots.

(A) Symmetrical funnel plot demonstrating that results from all studies are centered around common relative risk but those from small studies vary more from the mean than do those from large studies.

(B) Publication bias occurs when small negative studies (shown as open circles in [A]) are not published. This results in the asymmetrical funnel plot shown in (B).

(C) Smaller studies are often of lower methodological quality, producing biased results (open circles). This can also produce an asymmetrical funnel plot.

Source. Reprinted from Sterne JAC, Egger M, Smith GD: "Systematic Reviews in Health Care: Investigating and Dealing With Publication and Other Biases in Meta-Analysis." *BMJ* 323:101–105, 2001. Copyright 2001 BMJ Publishing Group. Used with permission.

Because small studies may be of lower methodological quality, they may be more likely to exaggerate the effects of a treatment (Kjaergard et al. 2001; Sterne et al. 2000, 2001). This situation, too, can produce an asymmetrical funnel plot, as demonstrated in Figure 6–2C.

Sensitivity Analysis

Questions often arise about whether a particular study or set of studies should be included in a meta-analysis because of differences in study design (e.g., patient population, treatment, co-intervention, outcome measure, duration, quality). In sensitivity analysis, the calculations are performed with and without a particular subset of studies to see whether this has any effect on the overall results (Egger and Smith 2001; Sutton et al. 2000).

A good example of the usefulness of sensitivity analysis comes from a systematic review of the effectiveness of antidepressants versus placebo in dysthymic disorder (Lima and Moncrieff 2002). Because the pre-1980 literature does not use the term *dysthymic disorder*, patients with this disorder were labeled as having *depressive neurosis*, *neurotic depression*, *depressive personality disorder*, and so forth. Lima and Moncrieff included studies of such patients but then used sensitivity analysis to show that the results were the same if the diagnosis were limited to *dysthymia* as they would be if these other studies were included.

Critical Appraisal Guide for Systematic Reviews

Even systematic reviews from respected sources can have methodological problems (Hopayian 2001; Olsen et al. 2001); therefore, all such reviews should be appraised by the reader. Several guides have been developed for the critical appraisal of systematic reviews (Badenoch and Heneghan 2002; Centre for Evidence-Based Medicine 2005; Freemantle and Geddes 1998; Geddes et al. 1998; Greenhalgh 2001; Oxman et al. 2002; Sackett et al. 2000; Seers 1999; Shea et al. 2001). One such guide appears in Table 6–3.

In 1996, a group of 30 clinical epidemiologists, clinicians, statisticians, editors, and researchers, concerned about the quality of meta-analyses, met to identify items they thought should be included in a checklist of standards. Whenever possible, checklist items were guided by research evidence suggesting that failure to adhere to the item proposed could lead to biased results. The conference resulted in the QUOROM (Quality Of Reporting Of Meta-analyses) statement, a checklist, and a flow diagram (Moher et al. 2000a). The checklist describes the group's preferred way to present the abstract, introduction, methods, results, and discussion sections of a report of a meta-analysis. The QUOROM statement, checklist, and flow diagram can be obtained online at www.consort-statement.org/index.aspx?o=1065.

Did the Review Address a Clearly Defined Issue?

The first step in planning a systematic review is formulating the question. The authors of the review should indicate the 4-part PICO question (Chapter 3) that they are trying to address. Unless the question itself is clear, the subsequent literature search likely will be unfocused. As a result, if the clinical question is not clear, the clinician should probably try to find another review (Seers 1999). However, a review also may be either too broad or too narrow (Oxman et al. 2002). For example, a systematic review of psychotherapy for mental illness is such a broad topic that it is unlikely that the reviewers could do it justice within the confines of even a book-length review. Conversely, a systematic review of a topic that is too narrow may not generate appropriate studies, and such a review may be of limited generalizability.

Did the Authors Select the Right Sort of Studies?

As noted in Table 4–1 in Chapter 4, some study designs are better than others for answering particular types of clinical questions. For example, as was discussed in Chapter 5, the RCT is the preferred study design for answering therapy questions. The authors of a review should specify their inclusion criteria for which types of studies were selected, and these criteria should be appropriate to the clinical question being asked (Seers 1999).

Were All Relevant Studies Included?

Several questions can be asked with regard to the studies included. How comprehensive was the search strategy? Did the authors search only MEDLINE, or did they also search other online data-

TABLE 6–3. Critical appraisal guide for systematic reviews

Did the review address a clearly defined issue?
Is the 4-part PICO question clearly identified?
Is the topic too broad or too narrow?
Did the authors select the right types of studies?
Are the inclusion criteria specified?
Do the authors specify the appropriate type of study to answer the question?
Were all relevant studies included?
How comprehensive was the search strategy?
Were appropriate electronic databases used?
Did the databases include non-English-language journals?
Did the authors go beyond electronic databases (e.g., personal contacts, manual searches)?
Was the quality of the studies addressed?
Were explicit criteria used?
Were two raters used, with a procedure for resolving differences?
Are the results similar from study to study? If not, was heterogeneity addressed?
Are the results clearly displayed (e.g., in a forest plot)?
Is there evidence of heterogeneity?
Are the reasons for the differences in study results discussed?
What are the overall results (with confidence intervals)?
What is the pooled effect measure, with confidence intervals?
Does it indicate that the two treatments are significantly different?
Can I apply the results to my patient?
Is my patient too different from those in the study?
What is the number needed to treat for my patient?
Is the treatment consistent with my patient's values and preferences?
Is the treatment feasible in my setting?

Note. PICO=patient/problem, intervention, comparison, outcome.
Source. Adapted from Centre for Evidence-Based Mental Health n.d.; Greenhalgh 2001; Oxman et al. 2002; Sackett et al. 2000; Seers 1999.

bases? Did the authors make other attempts to locate studies? For example, did they contact authors of published studies or manufacturers of drugs, check reference lists of published studies or prior reviews (i.e., snowballing), perform manual searches of journals or proceedings abstracts that are not abstracted in MEDLINE or similar databases, and check databases of clinical trials? If such measures are not taken, several sources of bias could influence the results, including publication bias and language bias (Table 6–2). Sampson et al. (2008) recently reviewed the criteria used to perform various searches. They found no clear consensus regarding optimum reporting of systematic review search methods, and commonly recommended items were not optimally reported.

Was the Quality of the Studies Addressed?

As was discussed earlier and in Chapter 5, evidence from less rigorously designed studies is more apt to be biased and misleading than is evidence from more rigorously designed studies. The authors of the

review should describe the criteria used for assessing quality, including the minimum quality needed for inclusion. For example, they may decide to include only RCTs with at least 80% follow-up and intention-to-treat analysis. Whatever the criteria, they should be explicit and decided in advance so that they are not influenced by study results (Silagy et al. 2002). In addition, the assessment of quality should generally involve two reviewers and a mechanism for resolving a difference between them (Glasziou et al. 2001; Oxman et al. 2002; Seers 1999). Reviews should list which studies were excluded from the analysis for quality reasons, and the reason for exclusion should be given.

Are the Results Similar From Study to Study?

Many systematic reviews present their results graphically in the form of forest plots. As previously discussed, forest plots can be used to assess whether the results are similar from study to study. The more formal statistical approach is to use a chi-square test for heterogeneity.

If heterogeneity is detected, an attempt should be made to assess the cause of the heterogeneity. Subgroup analysis may be appropriate if the rationale was identified beforehand. Sensitivity analysis also may indicate whether the inclusion of certain studies has a significant effect on the pooled measure of treatment effect.

What Are the Overall Results (With Confidence Intervals)?

The results of a meta-analysis generally will be presented as a forest diagram that includes a pooled effect measure and its CIs. For dichotomous outcome measures, this is an OR; for continuous measures, it is a standardized mean difference. If the CIs for the pooled OR include 1 or if the CIs for the standardized mean difference include 0, there is no statistically significant difference between the experimental and the control treatments.

Can the Clinician Apply the Results to His or Her Patient?

The final step in the appraisal process is to assess whether the clinician can apply the results to his or her patient. As with the discussion of a single therapy study in Chapter 5, several considerations enter into this assessment.

First, is the patient so different from those in the studies that the results do not apply? The question can be reframed in the following way: "Is the pathobiology of this patient so different from that of the study patients that the results cannot apply?" Typically, the answer to this question is "no," although there may be quantitative differences. To quantify the likelihood of a patient benefiting from a treatment, it is first necessary to convert the pooled estimate of the OR to NNT (McQuay and Moore 1998; Sackett et al. 2000) (see Appendix B).

Finally, the clinician must consider two additional questions before applying the treatment to his or her patient. First, is the treatment consistent with his or her patient's values and preferences? Second, is the treatment feasible in the clinician's setting? Assuming that the clinician has found a high-quality systematic review that has identified an effective treatment for the patient, the next steps in the EBM process (Table 2–1, Chapter 2) are to treat the patient and to evaluate the outcome.

References

Allen IE, Olkin I: Estimating time to conduct a meta-analysis from number of citations retrieved. JAMA 2892:634–635, 1999

Altman DG, Schulz KF: Statistics notes: concealing treatment allocation in randomised trials. BMJ 323:446–447, 2001

Antes G, Oxman AD: The Cochrane Collaboration in the 20th century, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 447–458

Badenoch D, Heneghan C: Evidence-Based Medicine Toolkit. London, BMJ Books, 2002

Balk EM, Bonis PAL, Moskowitz H, et al: Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. JAMA 287:2973–2982, 2002

Centre for Evidence-Based Medicine: Critical Appraisal Worksheet for Systematic Review of Therapy [Centre for Evidence-Based Medicine Web site]. 2005. Available at: http://www.cebm.net/index.aspx?o=1157. Accessed June 28, 2009

Centre for Evidence-Based Mental Health: Critical appraisal form for an overview [Centre for Evidence-Based Mental Health Web site]. Available at: http://cebmh.warne.ox.ac.uk/cebmh/downloads/education_resources/overview.pdf. Accessed June 28, 2009

Clarke M, Oxman AD (eds): Cochrane Reviewers' Handbook 4.1.6. [Cochrane Collaboration Web site]. April 2002. Available at: http://www.cochrane.org/cochrane/hbook.htm. Accessed September 5, 2002

Cochrane Collaboration: The Cochrane Collaboration leaflet [Cochrane Collaboration Web site]. Available at: http://www.cochrane.org/resources/leaflet.htm. September 1997. Accessed June 18, 2002

Cook DJ, Mulrow CD, Haynes RB: Synthesis of best evidence for clinical decisions, in Systematic Reviews: Synthesis of Best Evidence for Health Care Decisions. Edited by Mulrow C, Cook D. Philadelphia, PA, American College of Physicians, 1998, pp 5–12

Cook RJ, Sackett DL: The number needed to treat: a clinically useful measure of treatment effect. BMJ 310:452–454, 1995

Davies P, Boruch R: The Campbell Collaboration does for public policy what Cochrane does for public health. BMJ 323:294–295, 2001

Deeks JJ, Altman DG: Effect measures for meta-analyses of trials with binary outcomes, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 313–335

Deeks JJ, Altman DG, Bradburn MJ: Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 285–312

DerSimonian R, Laird N: Meta-analysis in clinical trials. Control Clin Trials 7:177–188, 1986

Dwan K, Altman DG, Arnaiz JA, et al: Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS ONE 3(8):e3081, 2008

Egger M, Smith GD: Meta-analysis: potentials and promise. BMJ 315:1371–1374, 1997

Egger M, Smith GD: Bias in location and selection of studies. BMJ 316:61–66, 1998

Egger M, Smith GD: Principles and procedures for systematic reviews, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 23–42

Egger M, Smith GD, Phillips AN: Meta-analysis: principles and procedures. BMJ 315:1533–1537, 1997a

Egger M, Smith GD, Schneider M, et al: Bias in meta-analysis detected by a simple, graphical test. BMJ 315:629–634, 1997b

Egger M, Dickersin K, Smith GD: Problems and limitations in conducting systematic reviews, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001a, pp 43–68

Egger M, Smith GD, Altman DG (eds): Systematic Reviews in Health Care: Meta-Analysis in Context. London, BMJ Books, 2001b

Egger M, Smith GD, O'Rourke K: Rationale, potentials, and promise of systematic reviews, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001c, pp 3–19

Freemantle N, Geddes J: Understanding and interpreting systematic reviews and meta-analyses, II: meta-analyses. Evid Based Ment Health 1:102–104, 1998

Furukawa TA, Guyatt GH, Griffith LE: Can we individualize the "number needed to treat"? An empirical study of summary effect measures in meta-analysis. Int J Epidemiol 31:72–76, 2002

Geddes J, Freemantle N, Streiner D, et al: Understanding and interpreting reviews and meta-analyses, I: rationale, search strategy, and describing results. Evid Based Ment Health 1:68–69, 1998

Glasziou P, Irwig L, Bain C, et al: Systematic Reviews in Health Care: A Practical Guide. Cambridge, UK, Cambridge University Press, 2001

Greenhalgh T: How to Read a Paper: The Basics of Evidence Based Medicine, 2nd Edition. London, BMJ Books, 2001

Guyatt G: Therapy and harm: why study results mislead—bias and random error, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 223–231

Helmer D, Savoie I, Green C, et al: Evidence-based practice: extending the search to find material for the systematic review. Bull Med Libr Assoc 89:346–352, 2001

Hopayian K: The need for caution in interpreting high quality systematic reviews. BMJ 323:681–684, 2001

Ioannidis JPA, Haidich A-B, Pappa M, et al: Comparison of evidence of treatment effects in randomized and nonrandomized studies. JAMA 286:821–830, 2001

Jüni P, Altman DG, Egger M: Assessing the quality of randomised controlled trials, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 87–108

Jüni P, Holenstein F, Sterne J, et al: Direction and impact of language bias in meta-analyses of controlled trials: empirical study. Int J Epidemiol 31:115–123, 2002

Kjaergard LL, Villumsen J, Gluud C: Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med 135:982–989, 2001

Kunz R, Oxman AD: The unpredictability paradox: review of hempirical comparisons of randomised and nonrandomised clinical trials. BMJ 317:1185–1190, 1998

Lacchetti C, Guyatt G: Therapy and validity: surprising results of randomized controlled trials, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 247–265

Lefebvre C, Clarke MJ: Identifying randomised trials, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 69–86

Lewis S, Clarke M: Forest plots: trying to see the wood and the trees. BMJ 322:1479–1480, 2001

Lima MS, Moncrieff S: Drugs versus placebo for dysthymia, in The Cochrane Library, Issue 3. Oxford, UK, Update Software, 2002

Mantel N, Haenszel W: Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 22:719–748, 1959

McAlister FA: Commentary: relative treatment effects are consistent across the spectrum of underlying risks…usually. Int J Epidemiol 31:76–77, 2002

McQuay HJ, Moore RA: Using numerical results from systematic reviews in clinical practice, in Systematic Reviews: Synthesis of Best Evidence for Health Care Decisions. Edited by Mulrow C, Cook D. Philadelphia, PA, American College of Physicians, 1998, pp 23–36

Moher D, Cook DJ, Eastwood S, et al: Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. QUOROM Group. Br J Surg 87:1448–1454, 2000a

Moher D, Klassen TP, Schulz KF, et al: What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol 5:964–972, 2000b

Moher D, Klassen TP, Jones AL, et al: Assessing the quality of reports of randomised trials included in meta-analyses: attitudes, practice, evidence and guides, in The Advanced Handbook of Methods in Evidence Based Healthcare. Edited by Stevens A, Abrams K, Brazier J, et al. London, Sage, 2001, pp 409–425

Montori V, Guyatt G: Summarizing the evidence: publication bias, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 529–538

Montori VM, Smieja M, Guyatt GH: Publication bias: a brief review for clinicians. Mayo Clin Proc 75:1284–1288, 2000

Montori V, Guyatt G, Oxman A, et al: Summarizing the evidence: fixed-effects and random-effects models, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 539–545

National Health Service Centre for Reviews and Dissemination [NHS CRD]: Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out commissioning reviews. CRD Report No 4, 2nd Edition [NHS CRD Web site]. March 2001. Available at: http://www.york.ac.uk/inst/crd/report4.htm. Accessed September 5, 2002

Olsen O, Middleton P, Ezzo J, et al: Quality of Cochrane reviews: assessment of sample from 1998. BMJ 323:829–832, 2001

Olson CM, Rennie D, Cook D, et al: Publication bias in editorial decision making. JAMA 287:2825–2828, 2002

Oxman A, Guyatt G, Cook D, et al: Summarizing the evidence, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 155–173

Petitti DB: Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine, 2nd Edition. New York, Oxford University Press, 2000

Pittler MH, Abbot NC, Harkness EF, et al: Location bias in controlled clinical trials of complementary/alternative therapies. J Clin Epidemiol 53:485–489, 2000

Robinson KA, Dickersin K: Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. Int J Epidemiol 31:150–153, 2002

Rochon PA, Bero LA, Bay AM, et al: Comparison of review articles published in peer-reviewed and throwaway journals. JAMA 287:2853–2856, 2002

Rosenthal R: The "file drawer problem" and tolerance for null results. Psychol Bull 86:638–641, 1979

Sackett DL, Haynes RB, Guyatt GH, et al: Clinical Epidemiology: A Basic Science for Clinical Medicine, 2nd Edition. Boston, MA, Little, Brown, 1991

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Sampson M, McGowan J, Tetzlaff J, et al: No consensus exists on search reporting methods for systematic reviews. J Clin Epidemiol 61:748–754, 2008

Schulz KF: Assessing allocation concealment and blinding in randomised controlled trials: why bother? Evid Based Ment Health 3:4–5, 2000

Schulz KF, Chalmers I, Hayes RJ, et al: Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 273:408–412, 1995

Seers K: Systematic review, in Evidence-Based Medicine: A Primer for Health Care Professionals. Edited by Dawes M, Davies P, Gray A, et al. Edinburgh, Churchill Livingstone, 1999, pp 85–100

Shea B, Dube C, Moher D: Assessing the quality of reports of systematic reviews: the QUORUM statement compared to other tools, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 122–139

Silagy CA, Middleton P, Hopewell S: Publishing protocols of systematic reviews: comparing what was done to what was planned. JAMA 287:2831–2834, 2002

Smith GD, Egger M: Going beyond the grand mean: subgroup analysis in meta-analysis of randomised trials, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 143–156

Song F, Eastwood A, Gilbody S, et al: Publication and related biases, in The Advanced Handbook of Methods in Evidence Based Healthcare. Edited by Stevens A, Abrams K, Brazier J, et al. London, Sage, 2001, pp 371–390

Stern JM, Simes RJ: Publication bias: evidence of delayed publication in a cohort study of clinical research projects. BMJ 315:640–645, 1997

Sterne JAC, Egger M: Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol 54:1046–1055, 2001

Sterne JAC, Gavaghan D, Egger M: Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol 53:1119–1129, 2000

Sterne JAC, Egger M, Smith GD: Systematic reviews in health care: investigating and dealing with publication and other biases in meta-analysis. BMJ 323:101–105, 2001

Sutton AJ, Abrams KR, Jones DR, et al: Methods for Meta-Analysis in Medical Research. New York, Wiley, 2000

Thompson SG: Why sources of heterogeneity in meta-analysis should be investigated. BMJ 309:1351–1355, 1994

Thompson SG: Why and how sources of heterogeneity should be investigated, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 157–175

Thornton A, Lee P: Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol 53:207–216, 2000

Tramer MR, Reynolds DJM, Moore RA, et al: Impact of covert duplicate publication on meta-analysis: a case study. BMJ 315:635–640, 1997

Turner EH, Matthews AM, Linardatos E, et al: Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med 358:252–260, 2008

West S, King V, Carey TS, et al: Systems to Rate the Strength of Scientific Evidence (Evidence Report/Technology Assessment, No 47). Rockville, MD, Agency for Healthcare Research and Quality, 2002

Yusuf S, Peto R, Lewis J, et al: Beta blockade during and after myocardial infarction: an overview of the randomized trials. Prog Cardiovasc Dis 27:335–371, 1985

This page intentionally left blank

7

Clinical Practice Guidelines

Gregory E. Gray, M.D., Ph.D.
C. Barr Taylor, M.D.

Clinical practice guidelines are generally defined as "systematically developed statements to assist practitioner decisions about appropriate health care for specific clinical circumstances" (Field and Lohr 1990, p. 1). Over the past two decades, there has been considerable interest and activity in developing these guidelines for a variety of clinical conditions, driven by concerns about variability in clinical practice, cost, quality, and legal liability (Birkmeyer 2001; Field and Lohr 1992; Greenhalgh 2001; Woolf et al. 1999b). We believe clinical practice guidelines are an invaluable source for improving patient care. Guidelines have many limitations, however, including the fact that most guidelines focus on a single disorder, that guidelines for treating one problem may conflict with aspects of guidelines for treating a comorbid problem in the same patient, and that many guidelines include recommendations that are not evidence based.

Role of Guidelines in Evidence-Based Practice

Most clinical practice guidelines represent an attempt to improve clinical care by focusing on effective evidence-based interventions. Although practice guidelines are sometimes equated with evidence-based medicine (EBM) (Grol 2001a), they should be viewed as distinct for several reasons (Gray 2002; Lipman 2000). First, as noted later in this chapter, not all practice guidelines are based on the best evidence, as derived from a systematic review of the medical literature (Browman 2001; Drake et al. 2001; Gray 2002; Greenhalgh 2001; Kahn et al. 1997; Woolf et al. 1999a, 1999b). All guidelines involve some degree of judgment and bias in their development (Browman 2001; Drake et al. 2001; Greenhalgh 2001; Greenhalgh and Peacock 2005), the extent of which is often unstated. In addition, clinical practice guidelines are often introduced as part of a "top-down" approach to changing clinician behavior, leading to clinician resistance (Haines and Jones 1994; Lipman 2000). Ideally, evidence-based psychiatric practice (EBPP) is a "bottom-up" approach in which clinicians make decisions on the basis of ability to search and appraise the medical literature.

From our standpoint, evidence-based clinical practice guidelines should play an important role in EBPP. Guidelines can serve as a very useful starting point for the practitioner to assess the major treatment principles that should be considered. Studies of practicing clinicians have found that clinicians often do not have the time to search and appraise the literature themselves but that they would welcome evidence-based guidelines when faced with clinical questions (McColl et al. 1998; Young and Ward 2001). In such cases, high-quality evidence-based guidelines can provide useful guidance (Feder et al. 1999; Grimshaw and Eccles 2001; Woolf et al. 1999b). In addition, such evidence-based guidelines generally either include or reference a systematic review of the relevant literature.

Sources of Guidelines

Clinical practice guidelines have been developed by a variety of organizations and are available for many DSM-IV-TR (American Psychiatric Association 2000) disorders. Some of the guidelines are evidence based; others are not. As described by Gray and Taylor in Chapter 4, a useful starting point in searching for evidence-based clinical practice guidelines is the National Guideline Clearinghouse (www.guideline.gov). A search function on the home page allows the user to find guidelines by disease, treatment, measures, and organization. Other resources include annotated bibliographies, expert commentaries, patient resources, syntheses, and comparison features. Other useful sources of guidelines are listed in Table 4–3.

American Psychiatric Association

The evidence-based guidelines most familiar to American psychiatrists are the ones developed by the American Psychiatric Association (APA) and represent a good starting point for the practitioner. These guidelines currently cover the major psychiatric disorders and are available both in print form (American Psychiatric Association 2006) and online (www.psychiatryonline.com). The APA guidelines are based on a systematic review of treatment options, with the use of expert opinion to synthesize the findings. One advantage of the APA guidelines is that they include a useful review of the relevant literature.

Expert Opinion

The APA, the Agency for Health Care Policy and Research (AHCPR), and the Agency for Healthcare Research and Quality (AHRQ) guidelines all involve expert opinion but with some quality assurance mechanisms, including standardized review of articles and databases, peer review and feedback, and public commentary. Also, several guidelines represent the opinion of experts. For instance, the *Expert Consensus Guidelines* (www.psychguides.com) provide a list of guidelines that can be purchased. In numerous instances, expert opinion on the treatment of medical conditions has been proved wrong by well-conducted research (Antman et al. 1992; Greenhalgh 2001; Mulrow 1994; Sackett et al. 1991); therefore, opinion does not necessarily provide the best guide to clinical practice. However, expert opinion can play a role in situations in which there is simply little or no evidence from well-conducted clinical trials. *Expert Review of Neurotherapeutics* (www.ingentaconnect.com/content/ftd/ern)provides expert reviews on the use of drugs and medicines in clinical neurology and neuropsychiatry. Coverage includes disease management, new medicines and drugs in neurology, therapeutic indications, diagnostics, medical treatment guidelines, and neurological diseases such as stroke, epilepsy, Alzheimer's disease, and Parkinson's disease. These are available by subscription. In Chapter 5, we noted the many potential biases in industry-sponsored trials. Expert opinion is also subject to many such biases.

Algorithms

Medication, and sometimes psychotherapy, algorithms are embedded in many guidelines. As mentioned in Chapter 4, several psychopharmacology algorithms are also available. Osser and Patterson have a Web site that lists many psychopharmacology algorithms (www.mhc.com/Algorithms). The project also lists the major guidelines and algorithm projects. The development and use of algorithms has been strongly influenced by the Texas Medication Algorithm Project (TMAP) started in 1996, which was designed to develop, implement, and evaluate a set of medication algorithms and an algorithm-driven treatment philosophy for major adult psychiatric disorders treated in the Texas public mental health sector. The ultimate goal of TMAP is to improve the quality of care and achieve the best possible patient outcomes for each dollar of resource expended. According to the developers, TMAP is a treatment philosophy for the medication management portion of care, consisting of 1) evidence-based, consensually agreed-on medication treatment algorithms; 2) clinical and technical support necessary to allow the clinician to implement the algorithm; 3) patient and family education programs that allow the patient to be an active partner in care; and 4) uniform documentation of care provided and resulting patient outcomes (Texas Medication Algorithm Project 2002). Currently, algorithms for schizophrenia, depression, and bipolar disorder are available (www.dshs.state.tx.us/mhprograms/tima.shtm). As with the *Expert Consensus Guidelines* (2002), the TMAP algorithms include an evidence base but are strongly influenced by expert consensus (Gilbert et al. 1998).

The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial has shown the benefit of using algorithms, particularly in the context of clearly defined and important clinical outcomes. The STAR*D trial is the largest study (*N*=4,041) ever conducted evaluating and comparing algorithmic treatment effectiveness in real-world patients experiencing a depressive episode as part of major depressive disorder (Trivedi et al. 2006). The STAR*D trial was supported by the National Institute of Mental Health (NIMH) and was implemented over a 5-year period. STAR*D was designed to assess effectiveness of treatments in generalizable samples and to ensure the delivery of adequate treatments, with the primary outcome measure being remission. The algorithm was designed to begin treatment with a selective serotonin reuptake inhibitor (SSRI). Citalopram was chosen as the initial drug (and SSRI) because of the absence of discontinuation symptoms, established safety in elderly and medically fragile patients, once-a-day dosing, few dose adjustment steps, and favorable drug-drug interaction profile (Trivedi et al. 2006). The protocol required that an adequate dose of citalopram be given for a sufficient time to ensure that an adequate treatment trial was conducted. A systematic but easily implemented approach to treatment, which the investigators referred to as *measurement-based care*, was developed.

Measurement-based care includes the routine measurement of symptoms and side effects at each treatment visit and the use of a treatment manual describing when and how to modify medication doses according to these measures. The details of how the algorithm can be implemented are provided by Trivedi and Kurian in Chapter 21, and the general algorithm has been described in many publications (Trivedi et al. 2006b). The study was designed to evaluate several issues, including identifying potential moderators of outcome. The overall results have been presented in many publications and should be familiar to most readers. The 8-week remission rates (28% for Hamilton Rating Scale for Depression; 33% for Quick Inventory of Depressive Symptomatology—Self-Report) were robust and similar to rates found in uncomplicated, nonchronic, symptomatic volunteers enrolled in placebo-controlled, randomized controlled trials with SSRIs (Agency for Health Care Policy and Research 1993b). These remission rates were better than those found in efficacy studies among patients with chronic depression (22%) (Keller et al. 2000). This trial sets a new standard of care for the treatment of depression in clinical practice.

Differences Among Guidelines

As indicated in the previous section, not all guidelines are based on the best available evidence. Some guidelines are based on expert consensus and are not truly evidence based (Berg et al. 1997; Browman 2001). Guidelines also differ considerably in comprehensiveness, format, frequency of review, and ease of use. Milner and Valenstein (2002) have provided a useful comparison of several of the guidelines for the treatment of schizophrenia that details many of these differences. Those produced by the APA, AHRQ, and the other organizations listed in Table 4–3 in Chapter 4 are generally of high quality, but any guideline should be assessed before its use.

Developing Evidence-Based Practice Guidelines

The development of evidence-based clinical practice guidelines is generally viewed as a 6-step process (Eccles et al. 2001; Shekelle et al. 1999a, 1999b) (Table 7–1).

Identify the Topic

The first step in the guideline development process is to identify and refine the subject of the guideline. Given the large number of diagnoses and the time needed to develop guidelines, some prioritization must occur. Often this step is based on considerations such as the prevalence or economic effects of a disorder (Berg et al. 1997; Cook et al. 1998; Shekelle et al. 1999a, 1999b). Furthermore, decisions must be made on the scope of the guideline (e.g., whether it should be restricted to pharmacological treatment or whether it should include psychosocial interventions, patient education, etc.). Some authors suggest creating a "causal pathway" that diagrams the links between steps in a diagnostic or treatment process and the potential outcomes (benefits or harms) that could occur (Berg et al. 1997; Shekelle et al. 1999a, 1999b).

Convene a Group

The group convened to develop the guideline should include individuals with expertise in statistics and epidemiology, as well as clinical expertise related

TABLE 7–1. Steps in developing evidence-based practice guidelines

1. Identify and refine the topic of the guideline
2. Convene an appropriate group
Typically 6–20 members
Requires both clinical and statistical expertise
Should be multidisciplinary
3. Gather and assess the evidence
Systematic review of literature
Hierarchy of evidence
4. Translate the evidence into recommendations
Some degree of clinical judgment always needed a) to weigh conflicting information and b) when there is little evidence
5. Use outside reviewers to review the recommendations
Should include users, as well as experts
Assess for validity and practicality
6. Update the guideline periodically

Source. Adapted from Berg et al. 1997; Eccles et al. 2001; Shekelle et al. 1999a.

to the condition or treatment that will be the topic of the guideline (Berg et al. 1997; Shekelle et al. 1999a, 1999b). Such groups typically have 6–20 members, with one member serving as the group leader to moderate discussions, often supported by a project management team (Shekelle et al. 1999a, 1999b). It is best to convene a multidisciplinary group that includes representation from all of the stakeholders involved in implementing the guideline (Cook et al. 1998; Shekelle et al. 1999a, 1999b). Concerns have been raised regarding the extent to which pharmaceutical company relationships may influence decisions by members of guideline development groups (Choudhry et al. 2002; Greenhalgh 2001).

Gather and Assess the Evidence

The third step is to gather and assess the evidence regarding the subject of the guideline. Preexisting systematic reviews can be helpful in this step, provided that the systematic reviews themselves are valid (Browman 2001; Cook et al. 1998; Eccles et al. 2001; Shekelle et al. 1999a, 1999b). If they are not, then the guideline development group must develop its own systematic review (see Gray, Chapter 6 in this volume).

In the process of assessing the evidence, different weight must be given to evidence from different study designs. As described in Chapter 4, hierarchies of evidence have been developed for this process (Badenoch and Heneghan 2002; Harbour and Miller 2001; Shekelle et al. 1999a, 1999b). For example, the National Institute for Health and Clinical Excellence (NICE) guidelines grade the level of evidence using the criteria provided in Table 7–2.

Translate the Evidence Into Recommendations

The fourth step is to translate the evidence into the recommendations that will make up the guideline. Clinical judgment is required in this step, both to weigh conflicting information and to make recommendations when little or no hard evidence exists (Harbour and Miller 2001; Shekelle et al. 1999a, 1999b). Various procedures have been developed to facilitate this process (Black et al. 2001; Kahn et al. 1997). Final recommendations should include an indication of the strength of evidence on which they are based (Harbour and Miller 2001; Pinsky and Deyo 2000; Shekelle et al. 1999a, 1999b). Again, the NICE guideline developers have developed a detailed procedure as to how to do this.

Forming and Grading the Statements and Recommendations

The standards and procedures developed by NICE are models for the field (National Institute for Health and Clinical Excellence 2007). As summarized from the guideline comparison search available through www.guideline.gov (and used in the comparison of guidelines discussed in the "Comparison of Guidelines" section later in this chapter), the NICE guidelines were developed as follows:

> The evidence tables and forest plots formed the basis for developing clinical statements and recommendations. For intervention studies, the statements were classified according to an accepted hierarchy of evidence. Recommendations were then graded A to C based on the level of associated evidence. In order to facilitate consistency in generating and drafting the clinical statements the guideline development group (GDG) utilised a statement decision tree. The flowchart was designed to assist with, but not re-

TABLE 7–2. Levels of evidence

I:	Evidence obtained from a single randomized controlled trial or a meta-analysis of randomized controlled trials
IIa:	Evidence obtained from at least one well-designed controlled study without randomization
IIb:	Evidence obtained from at least one well-designed quasi-experimental study
III:	Evidence obtained from well-designed nonexperimental descriptive studies, such as comparative studies, correlation studies, and case-control studies
IV:	Evidence obtained from expert committee reports or opinions and/or clinical experience of respected authorities

Source. Adapted from National Institute for Health and Clinical Excellence 2007.

> place, clinical judgment. Where a statistically significant summary statistic (effect size [ES]) was obtained (after controlling for heterogeneity), the GDG considered whether this finding was of clinical significance (i.e., likely to be of benefit to patients) taking into account the trial population, nature of the outcome, and size of the effect. On the basis of this consideration the ES was characterized as "clinically significant" or not. A further consideration was made about the strength of the evidence by examining the confidence interval (CI) surrounding the ES. For instance, for level I evidence, where the ES was judged to be clinically significant and had a CI entirely within a clinical relevant range, the result was characterised as "strong evidence" (S1).

The NICE guideline review process also includes the detailed methods used to answer a clinical question in the absence of appropriately designed high-quality research (National Institute for Health and Clinical Excellence 2007). The clinical recommendations are graded as follows:

- **Grade A**—At least one randomized controlled trial as part of a body of literature of overall good quality and consistency addressing the specific recommendation (evidence level I) without extrapolation
- **Grade B**—Well-conducted clinical studies but no randomized clinical trials on the topic of recommendation (evidence level II or III) or extrapolated from level I evidence
- **Grade C**—Expert committee reports or opinions and clinical experiences of respected authorities (evidence level IV) or extrapolated from level I or II evidence. This grading indicates that directly applicable clinical studies of good quality are absent or not readily available.

Obtain Outside Review

The fifth step in the guideline development process is external review to ensure that the guidelines are both valid and practical. This step should be done by potential users of the guidelines, as well as those with clinical and scientific expertise (Berg et al. 1997; Shekelle et al. 1999a, 1999b).

Update Guidelines Periodically

It is recognized that guidelines may become outdated as new diagnostic and treatment procedures are developed and as additional knowledge is acquired regarding the benefits and harms of existing procedures. On the basis of a review of published AHRQ guidelines, Shekelle et al. (2001) suggested that guidelines should be assessed for validity every 3 years, although the generalizability of this recommendation to other guidelines remains in doubt (Browman 2001). In any case, some strategy must exist for periodically evaluating new evidence and for updating guidelines as necessary (Browman 2001; Shekelle et al. 1999a, 1999b, 2001).

Critical Appraisal of Guidelines

Because of the variable nature of the guideline development process, with some guidelines linked more closely to the research evidence than others, it is necessary to critically appraise guidelines before following their recommendations (Feder et al. 1999). Several guidelines have been developed for the critical appraisal of clinical practice guidelines (Greenhalgh 2001; Grimshaw and Eccles 2001; Guyatt et al. 2002; Pinsky and Deyo 2000; Sackett et al. 2000; Snowball 1999). Table 7–3 presents one approach.

TABLE 7–3. Critical appraisal guide for a clinical practice guideline

Is the guideline valid?
Did the developers carry out a systematic review of the literature?
Were all relevant treatment options and outcomes considered?
Did the developers specify and make explicit the values associated with various outcomes?
Did the developers indicate the level of evidence (and sources) upon which each recommendation was based?
Is the guideline applicable to my practice?
Is the burden of illness too low to warrant implementation?
Are the beliefs of my patients incompatible with the guidelines?
Are the costs and other barriers of implementation too high?

Source. Adapted from Greenhalgh 2001; Guyatt et al. 2002; Sackett et al. 2000.

Is the Guideline Valid?

The first part of the appraisal concerns the validity of the clinical practice guidelines.

Did the Developers Carry Out a Systematic Review?

The development of a systematic review is what distinguishes evidence-based guidelines from opinion-based guidelines. The NICE methodology provides an excellent model for guideline development.

Were All Relevant Treatment Options and Outcomes Considered?

Several questions should be asked with regard to treatment options and outcome. For example, do the guidelines mention only pharmacotherapy, or do they include psychotherapy and other psychosocial interventions? Have the developers considered potential harms from the interventions, or have they considered only the benefits? It is more likely that these issues will be addressed if the group developing the guidelines is multidisciplinary rather than one limited by narrow expertise (Guyatt et al. 2002).

Did the Developers Specify Values Associated With Various Outcomes?

Reasonable people can make very different recommendations after considering the same evidence, depending on the value attached to different outcomes. In a review of atypical antipsychotics in the treatment of schizophrenia, Geddes et al. (2000) concluded that atypical and conventional antipsychotics were equal in effectiveness and tolerability. The authors then went on to recommend conventional antipsychotics as first-line drugs because of cost considerations. Kapur and Remington (2000) reached the opposite conclusion after reviewing the same data, giving more value to the differences in risk of extrapyramidal side effects and less value to other side effects and to economic considerations. In weighing benefits and risks, guideline developers should be clear as to how they arrive at a particular recommendation (Guyatt et al. 2002).

Did the Developers Indicate the Level of Evidence and Sources on Which Each Recommendation Was Based?

If guideline developers indicate the level of evidence and sources on which each recommendation is based, the user can determine which recommendations are based on strong evidence and which recommendations are based on opinion. Ideally, guidelines should be primarily based on research evidence rather than on opinion; however, guidelines typically have some recommendations based primarily on expert opinion, and these should be labeled as such.

Is the Guideline Applicable to My Practice?

Assuming that the guidelines are primarily evidence based (and that the recommendations that are opinion based are clearly indicated), the user still must decide whether to follow the recommendations.

Is the Burden of Illness Too Low to Warrant Implementation?

If a guideline involves screening for disease, is the risk high enough in your patient population to make this recommendation worthwhile? As discussed in

Chapter 9, screening low-risk populations produces far more false-positive than true-positive test results. Similarly, if the recommendation involves preventive measures, implementing them in a low-risk population may produce more harm than good (Sackett et al. 2000). Guidelines developed for use in tertiary care settings, where the severity of illness and degree of comorbidity are high, may not be applicable in primary care settings, where illness severity is lower (Greenhalgh 2001).

Are the Beliefs of My Patients Incompatible With the Guidelines?

If your patients prefer psychotherapy and the recommendation is for medications (or vice versa), the guideline will be difficult to implement.

Are the Costs and Other Barriers to Implementation Too High?

Perhaps the guideline recommends a type of therapy that is unavailable in your geographic area. Is it practical for someone to be trained in its application? Or perhaps the recommendation concerns a method of delivering services (e.g., assertive community treatment) that is not available in your setting. What are the costs and other barriers to implementation?

Comparison of Guidelines

Choosing one guideline over another can be a challenge. After reviewing and comparing guidelines, the clinician might decide to combine aspects of guidelines or even to include other recommendations. Presented with similar evidence, different guideline development groups might arrive at different conclusions. As an example, we compared the specific recommendations for the NICE and APA guidelines for the use of psychotherapies and medications in bulimia.

We began by using the comparison feature available on the National Guideline Clearinghouse Web site (www.guideline.gov) with the term *eating disorders*. To do this, we first searched the database to identify potentially useful guidelines and added two guidelines to our "collection": the National Collaborating Centre for Mental Health's *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders* (NGC:003550) and the American Psychiatric Association's *Practice Guideline for the Treatment of Patients With Eating Disorders*, American Psychiatric Association—Medical Specialty Society, 1993 (revised June 2006) (NGC:004987). We then went back to the home page to "Compare Checked Guidelines" and selected these two guidelines to compare. The program provided the basic characteristics of the two guidelines on several dimensions: guideline developers, sources of funding, composition of group that authored the guideline, conflicts of interest, guideline objectives, major outcomes considered, methods used to collect or select evidence, methods used to assess the quality and strength of the evidence, major recommendations, and so forth.

The NICE method seemed superior to the APA method, but both recommendations were worth considering. Unfortunately, the actual recommendations are difficult to compare. Both list more than 100 treatment recommendations (although many are specific to one disorder or another; e.g., anorexia or bulimia).

Another approach to comparing guidelines is to use the "Guideline Synthesis" function. For example, we used the "Guideline Synthesis" function to generate a list of available syntheses. We then selected "Management of Eating Disorders" from the list. The synthesis was conducted on three guidelines:

1. APA: *Practice Guideline for the Treatment of Patients With Eating Disorders*, 3rd Edition. Washington, DC, American Psychiatric Association
2. Finnish Medical Society Duodecim (FMSD): "Eating disorders among children and adolescents," in *EBM Guidelines: Evidence-Based Medicine* [Internet]. Helsinki, Finland, Wiley Interscience, John Wiley and Sons, 2007
3. National Collaborating Centre for Mental Health (NCCMH/NICE): *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders*. Leicester, UK, British Psychological Society, 2004.

The first part of the synthesis provides much of the same information as does the comparison feature but goes further in providing a written guideline comparison. For instance, the content comparison notes, "While FMSD focuses on anorexia and bulimia nervosa, APA and NCCMH/NICE also address atypical eating disorders such as binge eating disorder and

eating disorders not otherwise specified (EDNOS)." This implies that in searching for a guideline addressing atypical eating disorders, the FMSD would not be useful. More important, the synthesis examines areas of agreement and disagreement.

We then examined in more detail some specific recommendations related to the case of a young woman with bulimia presented in Chapter 18. The two guidelines would lead the clinician in somewhat different directions.

Table 7–4 lists the main psychotherapy recommendations for the two guidelines. The NICE guideline suggests starting with a self-help approach, which may not be feasible in private practice but might work well in a health maintenance organization. The NICE guideline also recommends cognitive-behavioral therapy (CBT) as the next choice. The APA notes that the evidence supports CBT as the most effective single intervention. However, the guidelines differ as to what to do if CBT fails or if the patient prefers another treatment. NICE states that interpersonal psychotherapy (IPT) should be considered as an alternative to CBT. The APA notes that IPT is effective in some cases but argues that CBT "is associated with more rapid remission of eating symptoms" and adds that "using psychodynamic interventions in conjunction with CBT and other psychotherapies may yield better global outcomes." A clinician curious about these different interpretations may want to go to the guideline to determine the database behind the recommendations, but this can be time-consuming. We searched PubMed with the terms *bulimia*, *psychotherapy* and with "limits of reviews," "summaries," and "meta-analyses." The most comprehensive review was by Wilson (2005), who stated:

> There is no evidence that an "integrated psychotherapy" is effective with BN [bulimia nervosa], let alone more effective than CBT alone. It has been suggested that complex cases characterized by comorbid personality disorders require a blend of CBT and psychodynamic psychotherapy.... No evidence supports this speculation. Controlled studies of the effectiveness of psychodynamic therapies are still conspicuously lacking.

The guidelines also differ on the use of medications (see Table 7–5). NICE recommends SSRIs but grades the recommendation as a C. APA rates the use of fluoxetine as I. NICE does not recommend other medications; APA notes that topiramate may be useful (III). We view these guidelines as sources of information.

Implementing Practice Guidelines and Algorithms

The reason that guidelines are usually developed is to attempt to improve the delivery of care within an organization, geographic area, or profession. The results of such attempts have been mixed (Bero et al. 1998; Davis and Taylor-Vaisey 1997; Greenhalgh 2001; Grimshaw and Russell 1993; Grol 2001b; Lipman 2000; National Health Service Centre for Reviews and Dissemination 1999; Owen et al. 2008; Oxman et al. 1995; Worrall et al. 1997). We mentioned in the beginning of the book that recent studies show that compared with treatment as usual, the use of algorithms and collaborative care approaches in the care of depressed patients enhances treatment outcomes by modifying practice procedures and treatment processes (Adli et al. 2006). Although implementation of STAR*D has not been compared with treatment without following STAR*D, the protocol is designed to achieve a clinically significant outcome. On the other hand, Owen et al. (2008) compared the effectiveness of a conceptually based, multicomponent "enhanced" strategy with a "basic" strategy for implementing antipsychotic management recommendations of Department of Veterans Affairs (VA) schizophrenia guidelines. Two VA medical centers in each of three Veterans Integrated Service Networks were randomly assigned to either a basic educational implementation strategy or the enhanced strategy, in which a trained nurse promoted provider guideline adherence and patient compliance. Patients with acute exacerbation of schizophrenia were enrolled and were assessed at baseline and 6 months, and their medical records were abstracted. The enhanced guideline implementation strategy increased addition of second-generation antipsychotic to first-generation antipsychotic therapy but did not significantly increase guideline-recommended switching from first-generation antipsychotic to second-generation antipsychotic monotherapy. Antipsychotic dosing was not significantly altered.

On the contrary, some organizations have found an improvement in quality of care and reduction of costs with the widespread implementation of guidelines (Mullaney 2005). Guidelines may be particu-

TABLE 7–4. A comparison of psychotherapy recommendations for treating bulimia for two guidelines

Area	NICE guideline	APA guideline
First step	**B**—As a possible first step, patients with bulimia nervosa should be encouraged to follow an evidence-based self-help program.	A variety of self-help and professionally guided self-help programs have been effective for some patients with bulimia nervosa **[I].**
Cognitive-behavioral therapy (CBT)	**A**—CBT for bulimia nervosa (CBT-BN), a specifically adapted form of CBT, should be offered to adults with bulimia nervosa. The course of treatment should be for 16–20 sessions over 4–5 months.	For treating acute episodes of bulimia nervosa in adults, the evidence strongly supports the value of CBT as the most effective single intervention **[I].**
If CBT does not work or patient prefers other treatments	**B**—When people with bulimia nervosa have not responded to or do not want CBT, other psychological treatments should be considered.	Some patients who do not respond initially to CBT may respond when switched to either interpersonal therapy (IPT) or fluoxetine **[II]** or other modes of treatment such as family and group psychotherapies **[III].**
Interpersonal psychotherapy (IPT)	**B**—IPT should be considered as an alternative to CBT, but patients should be informed that it takes 8–12 months to achieve results comparable with those of CBT.	Controlled trials also have shown the utility of IPT in some cases **[II].** In clinical practice, many practitioners combine elements of CBT, IPT, and other psychotherapeutic techniques. Compared with psychodynamic therapy or IPT, CBT is associated with more rapid remission of eating symptoms **[I],** but using psychodynamic interventions in conjunction with CBT and other psychotherapies may yield better global outcomes **[II].**

Note. APA=American Psychiatric Association; NICE=National Institute for Health and Clinical Excellence.
A=At least one randomized controlled trial; **B**=Well-conducted clinical studies but no randomized clinical trials. **[I]**=Recommended with substantial clinical confidence; **[II]**=Recommended with moderate clinical confidence; **[III]**=May be recommended on the basis of individual circumstances.

Source. American Psychiatric Association: *Practice Guideline for the Treatment of Patients With Eating Disorders*, 3rd Edition. Washington, DC, American Psychiatric Association, 2006; National Collaborating Centre for Mental Health (NCCMH/NICE): *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders*. Leicester, UK, British Psychological Society, 2004.

larly important when they recommend a treatment that is not widely used but might be associated with a better outcome. For instance, the algorithms for treating schizophrenia developed by TMAP, the International Psychopharmacology Algorithm Project, and the Psychopharmacology Algorithm Project at the Harvard South Shore Department of Psychiatry all recommend using clozapine after two adequate monotherapy trials of other antipsychotics in schizophrenia. The latest Clinical Antipsychotic Trials of Intervention Effectiveness data support this recommendation (McEvoy et al. 2006). Yet clinicians appear to prefer to try many additional monotherapy trials, various combinations of antipsychotics, and other polytherapy before starting clozapine. The reasons for not following this recommendation are not known but may be related to the fact that it is a more demanding treatment to implement for the physician and the patient. There may be concerns about adverse effects, additional time and effort to obtain consent, and the need for medical monitoring in addition to patient reluctance for some of the same reasons.

TABLE 7–5. Comparison of medication recommendations for treating bulimia for two guidelines

Area	NICE guideline	APA guideline
Initial medication	**C**—Selective serotonin reuptake inhibitors (SSRIs) (specifically fluoxetine) are the drugs of first choice for the treatment of bulimia nervosa.	To date, fluoxetine is the best studied of these SSRIs and is the only FDA-approved medication for bulimia nervosa. Sertraline is the only other SSRI that has been shown to be effective, as demonstrated in a small randomized controlled trial. In the absence of therapists qualified to treat bulimia nervosa with cognitive-behavioral therapy (CBT), fluoxetine is recommended as an initial treatment **[I].**
Other drugs	None are recommended.	Small controlled trials have confirmed the efficacy of the anticonvulsant medication topiramate, but because adverse reactions to this medication are common, it should be used only when other medications have proved ineffective **[III].**

Note. APA=American Psychiatric Association; FDA=U.S. Food and Drug Administration; NICE=National Institute for Health and Clinical Excellence.

C=Directly applicable clinical studies of good quality are absent or not readily available. **[I]**=Recommended with substantial clinical confidence; **[III]**=May be recommended on the basis of individual circumstances.

Source. Adapted from American Psychiatric Association: *Practice Guideline for the Treatment of Patients With Eating Disorders*, 3rd Edition. Washington, DC, American Psychiatric Association, 2006; National Collaborating Centre for Mental Health (NCCMH/NICE): *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa*, *Bulimia Nervosa and Related Eating Disorders*. Leicester, UK, , British Psychological Society, 2004.

In general, the passive dissemination of guidelines has little effect on clinical practice (Bero et al. 1998; Davis and Taylor-Vaisey 1997; Oxman et al. 1995) or on patient outcome (Farmer et al. 2008). Guidelines are more apt to be adopted if they take account of local circumstances, are disseminated by active educational interventions, and are implemented using patient-specific reminders (Bero et al. 1998; Feder et al. 1999; Greenhalgh 2001; National Health Service Centre for Reviews and Dissemination 1999). Multifaceted interventions tend to be more effective, but they are also more expensive (Bero et al. 1998; Davis and Taylor-Vaisey 1997; Greenhalgh 2001; Grol 1997; National Health Service Centre for Reviews and Dissemination 1999; Oxman et al. 1995). In the cases presented in Chapters 17–34, we discuss ways to use guidelines more effectively.

Improving the Use of Guidelines

Reluctance and resistance to using guidelines come from many sources. Curiously, clinicians may agree with the recommendations when they are presented to them separately (Cabana et al. 1999). Many clinicians probably have a working set of guidelines in their mind based on experience, training, and perhaps some updating of the literature.

With all these limitations, guidelines are still an important part of managing patients. As noted several times in this book, we review guidelines as a starting place. They are not a definitive source of patient care. However, clinicians should use guidelines as an important source of care options for their patients. The point is not to bring the patient to the guideline but to bring the guideline to the patient. The authors of some of the cases presented in this volume created checklists from the guidelines, and they used the checklists both to guide therapy and as a review to help them identify areas in which they could improve their practice.

Choosing Guidelines

Choosing which guideline or guidelines to follow for the complicated, comorbid cases we see in clinical practice is a challenge. Guidelines and medication algorithms are usually developed to treat most "single" disorders; few address comorbid problems. As illustrated in Chapter 18, a guideline for treating one disorder may be contradicted by a guideline for

treating a comorbid disorder in the same patient. Many clinicians may prefer to follow only one guideline, presumably the one relevant to the primary problem. Other clinicians may wish to combine algorithms and guidelines drawn from various sources, as illustrated in Chapters 18, 20, and 34 among others. As therapy progresses, new problems, not covered by the initial guidelines and algorithms, may emerge and require consideration of other evidence bases and approaches.

The treatment setting and nature of the clinician's practice also affect which guidelines are selected. A guideline developed for a 20- to 30-minute patient visit might be very different from a guideline developed for a 50-minute visit. A guideline for a 50-minute session involving a nondirective psychodynamic approach will be quite different from a guideline for a 50-minute directive, structured session. Clinicians can find guidelines and algorithms appropriate for their practice setting and situation, and it is important for clinicians to continue to update them as a result of their experience and new information.

Another major problem in following guidelines is that they tend to focus on treatment planning rather than on implementation. The latter is much more challenging for most clinicians. In Chapters 17–34, experienced clinicians present examples of how they have used guidelines, algorithms, and other sources of EBPP to provide care for the individuals they discuss. We believe the cases in this volume provide wonderful examples of how EBPP can improve patient care!

References

Adli M, Bauer M, Rush AJ: Algorithms and collaborative-care systems for depression: are they effective and why? A systematic review. Biol Psychiatry 59:1029–1038, 2006

Agency for Health Care Policy and Research: Depression in primary care, Vol 2: Treatment of major depression. Clinical Practice Guideline No 5. Rockville, MD, U.S. Department of Health and Human Services, 1993b

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

American Psychiatric Association: Practice Guidelines for the Treatment of Psychiatric Disorders, Compendium. Washington, DC, American Psychiatric Association, 2006

Antman EM, Lau J, Kupelnick B, et al: A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. JAMA 268:240–248, 1992

Badenoch D, Heneghan C: Evidence-Based Medicine Toolkit. London, BMJ Books, 2002

Berg AO, Atkins D, Tierney W: Clinical practice guidelines in practice and education. J Gen Intern Med 12 (suppl 2):S25–S33, 1997

Bero LA, Grilli R, Grimshaw JM, et al: Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 317:465–468, 1998

Birkmeyer JD: Primer on geographic variation in health care. Eff Clin Pract 4:232–233, 2001

Black N, Murphy M, Lamping D, et al: Consensus development methods and their use in creating clinical guidelines, in The Advanced Handbook of Methods in Evidence Based Healthcare. Edited by Stevens A, Abrams K, Brazier J, et al. London, Sage, 2001, pp 426–448

Browman GP: Development and aftercare of clinical guidelines: the balance between rigor and pragmatism. JAMA 286:1509–1511, 2001

Cabana MD, Rand CS, Powe NR, et al: Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 282:1458–1465, 1999

Choudhry NK, Stelfox HT, Detsky AS: Relationships between authors of clinical practice guidelines and the pharmaceutical industry. JAMA 287:612–617, 2002

Cook DJ, Greengold NL, Ellrodt AG, et al: The relation between systematic reviews and practice guidelines, in Systematic Reviews: Synthesis of Best Evidence for Health Care Decisions. Edited by Mulrow C, Cook D. Philadelphia, PA, American College of Physicians, 1998, pp 55–65

Davis DA, Taylor-Vaisey A: Translating guidelines into practice: a systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. CMAJ 157:408–416, 1997

Drake RE, Goldman HH, Leff HS, et al: Implementing evidence-based practices in routine mental health settings. Psychiatr Serv 52:179–182, 2001

Eccles M, Freemantle N, Mason J: Using systematic reviews in clinical guideline development, in Systematic Reviews in Health Care: Meta-Analysis in Context. Edited by Egger M, Smith GD, Altman DG. London, BMJ Books, 2001, pp 400–409

Expert Consensus Guidelines: Psychiatric treatment guidelines for the most difficult questions facing clinicians [Expert Consensus Guidelines Web site]. April 3, 2002. Available at: http://www.psychguides.com. Accessed September 13, 2002

Farmer AP, Légaré F, Turcot L, et al: Printed educational materials: effects on professional practice and health care outcomes. Cochrane Database Syst Rev (3):CD004398, 2008

Feder G, Eccles M, Grol R, et al: Using clinical guidelines. BMJ 318:728–730, 1999

Field MJ, Lohr KN (eds): Clinical Practice Guidelines: Directions for a New Program. Washington, DC, National Academy Press, 1990

Field MJ, Lohr KN (eds): Clinical Practice Guidelines: From Development to Use. Washington, DC, National Academy Press, 1992

Geddes J, Freemantle N, Harrison P, et al: Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. BMJ 321:1371–1376, 2000

Gilbert DA, Altshuler KZ, Rago WV, et al: Texas Medication Algorithm Project: definitions, rationale, and methods to develop medication algorithms. J Clin Psychiatry 59:345–351, 1998

Gray GE: Evidence-based medicine: an introduction for psychiatrists. J Psychiatr Pract 8:5–13, 2002

Greenhalgh T: How to Read a Paper: The Basics of Evidence-Based Medicine, 2nd Edition. London, BMJ Books, 2001

Greenhalgh T, Peacock R: Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ 331(7524):1064–1065, 2005

Grimshaw J, Eccles M: Identifying and using evidence-based guidelines in general practice, in Evidence-Based Practice in Primary Care, 2nd Edition. Edited by Silagy C, Haines A. London, BMJ Books, 2001, pp 120–134

Grimshaw JM, Russell IT: Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 342:1317–1322, 1993

Grol R: Personal paper: beliefs and evidence in changing clinical practice. BMJ 315:418–421, 1997

Grol R: Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. JAMA 286:2578–2585, 2001a

Grol R: Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care 39 (suppl 2):II-46–II-54, 2001b

Guyatt G, Hayward R, Richardson WS, et al: Moving from evidence to action, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 175–199

Haines A, Jones R: Implementing findings of research. BMJ 308:1488–1492, 1994

Harbour R, Miller J: A new system for grading recommendations in evidence based guidelines. BMJ 323:334–336, 2001

Kahn DA, Docherty JP, Carpenter D, et al: Consensus methods in practice guideline development: a review and description of a new method. Psychopharmacol Bull 33:631–639, 1997

Kapur S, Remington G: Atypical antipsychotics. BMJ 321: 1360–1361, 2000

Keller MB, McCullough JP, Klein DN, et al: A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. N Engl J Med 342:1462–1470, 2000

Lipman T: Evidence-based practice in general practice and primary care, in Evidence-Based Practice: A Critical Appraisal. Edited by Trinder L, Reynolds S. Oxford, UK, Blackwell Scientific, 2000, pp 35–65

McColl A, Smith H, White P, et al: General practitioners' perception of the route to evidence based medicine: a questionnaire survey. BMJ 316:361–365, 1998

McEvoy JP, Lieberman JA, Stroup TS, et al: Effectiveness of clozapine versus olanzapine, quetiapine, and risperidone in patients with chronic schizophrenia who did not respond to prior atypical antipsychotic treatment. Am J Psychiatry 163:600–610, 2006

Milner KK, Valenstein M: A comparison of guidelines for the treatment of schizophrenia. Psychiatr Serv 53:888–890, 2002

Mullaney TJ: Doctors wielding data. Business Week, November 21, 2005, pp 94, 98

Mulrow CD: Systematic reviews: rationale for systematic reviews. BMJ 309:597–599, 1994

National Health Service Centre for Reviews and Dissemination: Getting evidence into practice. Eff Health Care 5(1):1–16, 1999

National Institute for Health and Clinical Excellence: How NICE clinical guidelines are developed: an overview for stakeholders, the public and NHS. London, National Health Service, 2007

Owen RR, Hudson T, Thrush C, et al: The effectiveness of guideline implementation strategies on improving antipsychotic medication management for schizophrenia. Med Care 46:686–691, 2008

Oxman AD, Thomson MA, Davis DA, et al: No magic bullet: a systematic review of 102 trials of interventions to improve professional practice. CMAJ 153:1423–1431, 1995

Pinsky LE, Deyo RA: Clinical guidelines: a strategy for translating evidence into practice, in Evidence-Based Clinical Practice: Concepts and Approaches. Edited by Geyman JP, Deyo RA, Ramsey SD. Boston, MA, Butterworth-Heinemann, 2000, pp 119–123

Sackett DL, Haynes RB, Guyatt GH, et al: Clinical Epidemiology: A Basic Science for Clinical Medicine, 2nd Edition. Boston, MA, Little, Brown, 1991

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Shekelle PG, Woolf SH, Eccles M, et al: Developing clinical guidelines. West J Med 170:348–351, 1999a

Shekelle PG, Woolf SH, Eccles M, et al: Developing guidelines. BMJ 318:593–596, 1999b

Shekelle PG, Ortiz E, Rhodes S, et al: Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 286:1461–1467, 2001

Snowball R: Critical appraisal of clinical guidelines, in Evidence-Based Practice: A Primer for Health Care Professionals. Edited by Dawes M, Davies P, Gray A, et al. Edinburgh, Churchill Livingstone, 1999, pp 127–131

Texas Medication Algorithm Project: TMAP—a collaborative effort [Texas Medication Algorithm Project Web site]. April 3, 2002. Available at: http://www.dshs.state.tx.us/mhprograms/TIMA.shtm. Accessed July 1, 2009

Trivedi MH, Rush AJ, Wisniewski SR, et al: Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry 163:28–40, 2006

Wilson GT: Psychological treatment of eating disorders. Annu Rev Clin Psychol 1:439–465, 2005

Woolf SH, Grol R, Hutchinson A, et al: Clinical guidelines: international overview [BMJ Web site]. February 20, 1999a. Available at: http://bmj.com/cgi/content/full/318/7182/527/DC1/1. Accessed September 24, 2002

Woolf SH, Grol R, Hutchinson A, et al: Potential benefits, limitations, and harms of clinical guidelines. BMJ 318:527–530, 1999b

Worrall G, Chaulk P, Freake D: The effects of clinical practice guidelines on patient outcomes in primary care: a systematic review. CMAJ 156:1705–1712, 1997

Young JM, Ward JE: Evidence-based medicine in general practice: beliefs and barriers among Australian GPs. J Eval Clin Pract 7:201–210, 2001

This page intentionally left blank

8

Measurement

C. Barr Taylor, M.D.

Measuring and monitoring patient progress are key to practicing evidence-based psychiatry. For many decades, researchers, policy makers, insurers, and some clinicians have noted the need to focus as much on outcomes as on process (Ellwood 1988; McGrath and Tempier 2003). Recently, the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) investigators have argued that "measurement" should be the paradigm for psychiatric practice (Trivedi and Daly 2007; Trivedi et al. 2007). Yet few clinicians routinely monitor outcomes unless compelled to do so. Why? The main reason is probably that they have little incentive to do so and some good reasons not to. In a busy practice, the added paperwork could be seen as a needless addition to an already pressured schedule. Even simple assessment forms and instruments need to handed out, scored, stored, and interpreted, contributing to increased "charting burden." Some clinicians believe that paper assessment, at least beyond the initial sessions, can interfere with the therapeutic relationship—a piece of paper is substituted for a more meaningful interaction. The clinician might argue that if patients seem to be improving and seem satisfied with their care, then they should not make changes. Ultimately, the clinician must decide that the effort results in better patient care.

Fortunately, obtaining and using assessment instruments have become easier in recent years. First, numerous assessment instruments can be readily obtained or purchased online. Many instruments relevant to psychiatric practice are available on the CD-ROM that accompanies the American Psychiatric Publishing *Handbook of Psychiatric Measures*, 2nd Edition (Rush et al. 2008). Copyright protection and other issues may limit the use of some standardized assessment instruments, but alternative public domain instruments usually can be found to measure outcomes. Second, practice management software systems are now available that make assessment and record keeping easier, as discussed at the end of this chapter. Some of the cases illustrate how measurement can be incorporated into clinical practice. Third, many guidelines and algorithms use measurement to guide decision making, and adherence to such guidelines and algorithms may improve practice. Fourth, evidence-based medicine (EBM) and evidence-based psychiatric practice (EBPP) are based on measurement.

In this chapter, I discuss "what to measure," review how one can consider the reliability and validity of a measure, present examples of some "nonstandard" measures, and provide examples of how process and outcome measures have been implemented in several practice settings.

Which Outcomes to Measure

Several factors need to be considered in deciding what to measure.[1] The type of measures may be

[1] Instruments used to assess baseline function, psychological process, and other phenomena may be different from instruments used to assess outcome. In this section, we focus on outcome assessment instruments.

determined by the practice setting or by the insurer. For instance, the Department of Veterans Affairs (VA) is adopting outcome measures for a variety of problems. Some insurers ask providers to complete periodic assessments. Domains of outcome measurement to consider include symptomatology, patient functioning, quality of life, patient satisfaction, and even cost and cost-benefit ratio. We recommend a smorgasbord approach. If a clinician is using a guideline or algorithm that has a measure to guide treatment decisions, then that measure should be adopted. For instance, those who are following the STAR*D guidelines would want to use a version of the Quick Inventory of Depressive Symptomatology (QIDS), and those using the Beck model of cognitive-behavioral therapy (CBT) would want to use the Beck Depression Inventory–II (BDI-II) and also, perhaps, the QIDS because it has defined outcomes for remission and relapse.

If a clinician has decided to follow practice recommendations from the evidence base for a particular problem and the analyses have used an available instrument, that would be a logical choice as a measure. For example, the Hamilton Rating Scale for Depression (Ham-D) is a widely used measure for depression outcome studies and might serve as a good reference in helping clinicians determine how well their patients are doing against published standards.

The clinicians' values and interests should also guide what domains they assess. For instance, one might want to include a quality-of-life measure (see Chapter 30). Second, one should consider measuring behaviors, attitudes, and issues of relevance to individual patients but not necessarily covered in standard instruments. Table 16–1 in Chapter 16 lists the measures used by the therapists in the case presentations. Some of the measures are standardized outcomes, with extensive psychometric data (e.g., the Posttraumatic Stress Disorder [PTSD] Checklist, the QIDS–Self-Report, the Yale-Brown Obsessive Compulsive Scale). Others represent self-reports of behaviors (e.g., binge and purge frequency, alcohol use), relationships, symptoms (e.g., dizziness, panic attacks, trouble swallowing), or medical parameters (e.g., weight).

The *Handbook of Psychiatric Measures*, 2nd Edition (Rush et al. 2008), has provided an invaluable resource of standard measures for clinicians. Table 8–1 lists the categories of measures available in this publication. The text of the *Handbook* includes a description of each measure (e.g., number of items), the goals of the measure, practical issues (how long does it take to administer, where permission can be obtained), the psychometric properties of the instrument (reliability, validity), and the clinical utility (what useful information it provides).

Constructing a Measure

Given the importance of measures, it is worthwhile to consider how they are developed. Measures can be constructed in several ways. Many of the classic psychiatric measurement skills were developed by a researcher simply trying to identify the main features of a disorder and then finding a way to rate the frequency and intensity of these features. For instance, when Hamilton decided to develop a measure for depression in the 1950s, he presumably began by listing the main features: depressed mood, feelings of guilt, suicidality, insomnia, trouble with work and activities, psychomotor retardation, agitation, mental and somatic anxiety, gastrointestinal and other somatic symptoms, hypochondriasis, change in weight, change in insight, diurnal variation of mood, depersonalization and derealization, paranoia, and obsessionality. He then developed a system to rate each of these items. Hamilton probably erred in mixing nominal and ordinal items and in assuming that each item was more or less equally as important as any other item (Babgy et al. 2004).

Beck developed his inventory (Beck et al. (1961; perhaps the most widely used self-report depression instrument in psychiatry) by first identifying symptoms and attitudes that were common in depressed patients but rare in nondepressed patients. These ideas were systematically consolidated into 21 symptoms and attitudes that could be rated from 0 to 3 in intensity. The items included mood, pessimism, sense of failure, lack of satisfaction, guilt, sense of punishment, sense of self-dislike, self-accusation, suicide wishes, crying, irritability, social withdrawal, fatigability, and loss of appetite. Seemingly, individuals who score higher are more depressed. Both the Ham-D and the BDI-II have undergone some revision, but the core measures remain much the same.

Both Hamilton and Beck used clinical observation to select items to measure. Some other approaches for scale construction are available. Another ap-

TABLE 8–1. Categories of measures available in *Handbook of Psychiatric Measures*

Measures are provided for:
Diagnosis in adults
General psychiatric symptoms
Mental health status, functioning, and disabilities
General health status, functioning, and disabilities
Quality of life
Adverse effects
Patient perceptions of care
Stress and life events
Family risk factors
Suicide risk
Diagnosis and screening in children and adolescents
Disorders usually first diagnosed in infancy, childhood, or adolescence
Child and adolescent functional status
Personality traits and defense mechanisms
Delirium and behavioral symptoms of cognition
Aggression
Cognitive, substance use, psychotic, mood, anxiety, somatoform, factitious and malingering, dissociative, sexual dysfunction, eating, sleep, impulse-control, and personality disorders

Source. Rush AJ Jr, First MB, Blacker D (eds): *Handbook of Psychiatric Measures*, 2nd Edition. Washington, DC, American Psychiatric Publishing, 2008.

proach would be to develop a rating scale based on the standard items of some diagnostic system, such as DSM-IV-TR (American Psychiatric Association 2000). This was the approach used in developing the QIDS, which was used in the STAR*D project (Rush et al. 2003). Another approach is to ask many questions of a population of interest and then to use statistical methods to identify those items that best distinguish one population (such as depressed individuals) from another (nondepressed individuals). The items then can be used to form a scale. We used this approach several years ago when we wanted to develop a measure to identify students at risk for developing an eating disorder (Killen et al. 1993). We gave various standard eating disorder attitude and behavior scales to 967 sixth- and seventh-grade girls. A principal components analysis was conducted to reduce redundancy to find a set of independent variables. The result was a set of five items measuring worry about weight and body shape, weight gain, dieting, importance of weight, and feeling fat.

Once items are identified, the next issue in scale development is to determine how to rate them. Items can be rated in several ways. For instance, they can be rated in terms of frequency (never to very often) or intensity (not at all to very) across dimensions (see the Ham-D, item 5, for an example of this). Books and articles on scale development are available for the interested reader (e.g., DeVellis 2003). Before giving forms to a patient, it is a good idea for the clinician to fill out a copy to get a sense of exactly what is being asked.

Psychometric Properties of Scales

Scientific studies favor standard measures with known and acceptable reliability and validity.

Reliability

In statistics, *reliability* is the consistency of a set of measurements or a measuring instrument. Three types of reliability are reported, depending on the purpose of the measure and how it is administered. Clinical assessment instruments report the internal reliability (or consistency) of individual items, the test-retest reliability of the total score and/or items in the measurement instrument, and interrater reliability of the instrument. Reliability is inversely related to random error; that is, high reliability scores make one more confident that the scale is measuring what it is supposed to.

Internal consistency represents how well the items of a scale measure a single dimension of an underlying construct. Because individuals may use different words to describe some aspect of their functioning, a scale that assesses a dimension with several questions might allow for more people to be able to report more accurately. Some measurement instruments include several scales, each of which might have "good internal consistency." Good internal consistency of a scale may actually limit its usefulness for some purposes. For instance, an instrument to measure mania may assess both psychotic and neurovegetative symptoms. Because these symptoms are not strongly correlated in bipolar patients, the measure would

have poor internal consistency, although it might accurately measure mania from a clinical standpoint.

Interrater reliability is a measure of agreement between two or more observers evaluating the same series of subjects and using the same information. Interrater reliability is best measured with the κ statistic, which is a measure of agreement corrected or changed (Fleiss 1981). For categorical measures, a κ value of 1 indicates perfect agreement; a κ of 0 indicates no agreement. In clinical trials, κ's of greater than 0.6–0.7 are considered adequate, and a κ greater than 0.8 is considered excellent (Landis and Koch 1977). The clinician should be wary of studies with a κ less than 0.5 because it suggests that two raters cannot agree on what they are judging. Reliability for continuous measures is more complicated. Suppose a rater is asked to determine whether a patient does not look at all sad, looks a little sad, looks somewhat sad, looks very sad, or looks very, very sad. It is harder for raters to agree on such levels, and statistical methods have been developed to account for this. Interrater reliability is extremely important for research studies.

Test-retest reliability measures whether the same observers or individuals completing a measure answer questions the same way from one time to the next, assuming there is no intervening change. Typically, the same test is repeated a few days or weeks apart. For retest reliability of scale items, a Pearson *r* greater than 0.70 is considered acceptable (Anastasia and Urbina 1997).

Validity

Reliability does not prove that a scale is valid. *Validity* refers to how well an instrument measures what it is supposed to. Types of validity of psychiatric rating scales include content, convergent, discriminant, factorial, predictive, and cultural validity.

Content validity is a nonstatistical type of validity. The basic issue is whether the content of the scale seems to cover a representative sample of the phenomenon to be measured. Content validity is assessed by examining scale items to determine correspondence with known features of a syndrome. Some of the very standard assessment instruments, such as the Ham-D, which was developed before the advent of the DSM-IV (American Psychiatric Association 1994) system, have been criticized for not including content relevant to the new diagnostic system, as discussed later in this chapter (Bagby et al. 2004) (see section "Reliability and Validity of the Hamilton Rating Scale for Depression"). The clinician also should determine whether the items measure areas that are expected to change with therapy.

Concurrent validity refers to the degree to which the measure correlates with other measures of the same construct measured at the same time. Concurrent validity is adequate when a scale shows Pearson *r* values of at least 0.50 in correlation with other measures of the same syndrome.

Other types of validity are important, depending on the purpose of the scale. *Discriminant validity* refers to how well the measure distinguishes groups differing in their diagnostic status. *Predictive validity* refers to how well the measure can predict the onset of symptom(s). *Cultural validity* refers to how well the measure works in various populations.

Sensitivity to Change

For outcome studies, measurements should be used that are likely to be able to detect change in the domain assessed. In general, sensitivity to change is reflected in the effect size (see Gray, Chapter 6 in this volume). Within a treatment group, effect size is often calculated as the difference between the mean score at time 1 and the mean score at time 2 for each subscale, divided by the standard deviation at time 1 or over time. Measures that have large standard deviations relative to the mean require either very strong interventions to show effects or very large sample sizes. An effect size of approximately 0.2 is generally somewhat arbitrarily considered to be small, one of 0.5 indicates moderate differences, and that of 0.8 or higher indicates large differences (Cohen 1988). Of course, an instrument may be very sensitive to change, but the change itself may be clinically meaningless. For this reason, some researchers argue that categorical measures (e.g., how many patients depressed at baseline are no longer depressed at posttreatment or follow-up on the basis of some criteria such as DSM-IV-TR) are more meaningful than ordinal measures.

A second approach is to define an outcome on a measure as indicated improvement; for instance, to expect at least a 50% reduction in a baseline measure or a score below a certain level on a scale such as the Ham-D. STAR*D provides guidelines for what constitutes remission or improvement (see Chapter 21). A third approach is to estimate how many patients fit

into the "nondysfunctional" range of the outcome measure or measures. For example, in a study of panic attack interventions, Newman et al. (1997) and later Kenardy et al. (2003) used norms and test-retest reliability estimates from previous studies to determine the dysfunctional/functional range. A successful outcome would be that, following treatment, a participant was no longer in the dysfunctional range, and his or her change from pre- to posttest exceeded measurement error. In addition, Jacobson and Truax (1991) required subjects to meet both a reliable change index and a functional recovery criterion, which was defined as being panic attack free. A fourth approach is to determine what would be considered a minimal clinically important difference, defined as being the smallest difference in score in the domain of interest that patients perceive as beneficial and that would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management (Jaeschke et al. 1989). This is a complicated issue, and some doubt exists that it is possible to arrive at a minimal clinically important difference (Beaton et al. 2002).

Reliability and Validity of the Hamilton Rating Scale for Depression

The Ham-D is the most frequently used measure for depression in psychopharmacology studies (Williams 2001), and changes in its score are used to judge the effectiveness of a medication. The Ham-D was developed in the late 1950s to assess the effectiveness of the first generation of antidepressants and was originally published in 1960 (Hamilton 1960). How reliable and valid is the Ham-D? To answer this question, Bagby et al. (2004) reviewed the psychometric properties of the 17-item Ham-D, perhaps the most widely used version. They used MEDLINE to identify all studies published during the period from January 1980 to May 2003 that addressed the psychometric properties of the Ham-D. Seventy articles met their selection criteria and were categorized into three groups on the basis of the major psychometric property examined—reliability, item response, or validity.

Reliability

The internal reliability of the Ham-D ranged from 0.46 to 0.97. However, 10 of 13 studies reported internal reliabilities greater than 0.70. All of the single items of the test also showed good reliability except for "loss of insight," which was quite variable. Overall, the Ham-D has good internal reliability, suggesting that the items measure one construct.

Interrater reliabilities were quite high: total Ham-D interrater reliabilities ranged from 0.82 to 0.98, and the intraclass *r* ranged from 0.46 to 0.99. However, a large range was seen on the individual items. This suggests that two individuals who were trained to administer the Ham-D and rated a depressed individual came up with similar scores, although they varied as to how negatively they rated one item to the next.

The overall retest reliability of the Ham-D was high, ranging from 0.81 to 0.98. However, the test-retest reliability of individual items ranged from 0.00 to 0.85. This suggests that over time with no intervening treatment, an individual is rated as having the same overall extent of depression from one time to the next. However, changes in particular items are much less reliable. Bagby et al. (2004) criticized how the scale was constructed, noting, for instance, that some items require ratings that may even cross dimensions. For example, the question of hypochondriasis is rated from self-absorption to delusions. Are hypochondriacal delusions on the same continuum as hypochondriacal symptoms?

Validity

Content Validity

Bagby et al. (2004) also questioned the content validity of the Ham-D, noting that the items were generated well before DSM-IV, and asked if the Ham-D accurately reflects depression as defined by DSM-IV. Important features of DSM-IV depression are buried within more complex items in the Ham-D and sometimes are not captured at all. The interested reader may want to examine the Depression Interview and Structured Hamilton (DISH), a structured interview developed for the National Heart, Lung, and Blood Institute's Enhancing Recovery in Coronary Heart Disease trial that combines the Ham-D and the clinical items of DSM-IV (Freedland et al. 2002).

Convergent Validity

The Ham-D has good *convergent validity* with many other scales given at the same time. In arguing, how-

ever, that the Ham-D does not reflect DSM-IV items, poor convergent validity between the Ham-D and the Structured Clinical Interview for DSM-IV was reported in a Turkish study. Language/translation issues might have been a factor in this result.

Discriminant Validity

The Ham-D, at cut-off points of greater than 1.0, had reasonable sensitivity (most studies reported sensitivities in the 0.7–0.8 range) and specificity (0.75–1.0).

Predictive Validity and Sensitivity to Change

Several very large meta-analyses involving thousands of patients and a variety of treatments have shown that the Ham-D is more sensitive to change than is the BDI, a widely used self-report instrument (Edwards et al. 1984; Lambert et al. 1986). Such studies also suggest that the instrument is sensitive to change.

Factorial Validity

Overall, the results from 15 studies suggested that the Ham-D is not unidimensional because separate sets of items seem to represent general depression and insomnia factors; however, the exact structure of its multidimensionality remains unclear.

Bagby et al. (2004) concluded that the Ham-D has adequate internal, interrater, and retest reliability and good convergent, discriminant, and predictive validity. They noted that the same is not true for individual items. For this reason, they recommended that the Ham-D should be replaced by other instruments and noted that the Inventory of Depressive Symptomatology (Rush et al. 1986) and the Montgomery-Åsberg Depression Rating Scale (Montgomery and Åsberg 1979), designed to address the limitations of the Ham-D, represent two potential replacement alternatives. The QIDS is obviously another one.

Critics of the Bagby et al. (2004) article noted that their analyses included studies in which the Ham-D was not used as intended, to measure change in depression, but to assess nondepressed populations (Corruble and Hardy 2005). A scale designed for one population may have excellent psychometric properties but do poorly in another population. For instance, in a "normal" population, in which very low scores on most items are expected, any small change may generate poor agreement. Carroll et al. (2005), also critical of Bagby et al. (2004), argued that the Ham-D may not be perfect but that it captures the many different ways that depression presents and has proved sensitive to change, thus providing a useful outcome to determine the effectiveness of medications.

Changing a measure that has been the gold standard for evidence-based trials creates many problems. Whereas statistical approaches can "adjust" between measures, the same measure used in the same way for the same presenting problem allows for easy comparison. Although this is not a compelling reason to continue using the Ham-D, a PubMed search (July 20, 2008) identified hundreds of citations since Bagby et al. (2004) recommended that it be abandoned. It may not be perfect, but it endures. From a practical standpoint, a clinician could use the Ham-D total score to measure improvement but be less confident that changes in the individual items are meaningful. I favor use of the QIDS-SR_{16} because it is tied into a measurement outcome–based protocol.

Practical Issues in Finding and Using Measures

An instrument with excellent psychometric properties may not be useful in clinical practice for several reasons. Some self-report measures include items that are reverse scored so that the clinician needs to have a scoring sheet or to have the form scored by a vendor or through software. The purpose of the reverse scoring is to help ensure that the person filling out the form reads each item and does not simply choose one end of the scale. Other assessments can be very lengthy and require a long time to complete. Many forms are also copyrighted. Simply because a form is available on the Internet does not mean that it is in the public domain and can be used for free or without permission. It is illegal and unethical to use a copyrighted test without buying it or receiving permission to use it from the owner of the copyright. American Psychiatric Publishing provides contact information for the use of the forms included on the CD-ROM accompanying the *Handbook of Psychiatric Measures* (Rush et al. 2008).

One may be surprised to learn that forms that have been used for some time are actually copyrighted. For instance, the Mini-Mental State Exam (MMSE) is now copyrighted and must be purchased. As an example of what forms cost, a package of 50

MMSE test forms costs $60. The American Psychiatric Publishing *Handbook of Psychiatric Measures*, 2nd Edition (Rush et al. 2008), includes a "Practical Issues" section for most measures that mentions how long it takes to complete the form, whether the form is copyrighted, and how it can be obtained.

Developing Your Own Assessment Measures

Some therapy goals can be measured without relying on standardized instruments. For instance, many behaviors, such as frequency of binge eating or drinking or the number of panic attacks experienced, can be assessed by self-report. The actual number of events can be counted, or the frequency can be described more generally. For instance, a patient may be experiencing nausea for psychological reasons. To get a sense of how the patient is doing, the clinician asks him to rate his nausea over the last week on the basis of how often he experienced it (Figure 8–1). The clinician also could ask the patient to rate it by severity (Figure 8–1). Table 16–1 in Chapter 16 of this volume lists cases in which the authors used a variety of nonstandardized measures to help guide their therapy.

A more complicated, and perhaps more interesting, issue is to determine how to "measure" complex phenomena specific to a particular patient. For instance, one might want to monitor the patient's "insight," satisfaction with relationships, confidence in conducting a job interview, or any number of the many phenomena that we address in therapy. Sometimes, I use a very simple satisfaction measure to assess progress. Figure 8–2 provides an example of a simple scale to measure confidence and satisfaction with dating in a patient with relationship issues (see also Chapter 18). The figure also uses a scale to rate the patient's confidence to use parenting skills. Such measures help remind me what the therapy is focused on and can be used to determine whether progress is occurring.

When I began to focus more on evidence-based practice, I decided to implement a very simple system that could work in most outpatient practice settings. I developed a weekly check-in form that evaluates mood (depression and anxiety) and includes measures specific to the patient. I also created a blank form that can be used to measure two target symptoms (per page) by frequency and severity, as appropriate. In the waiting room or at the beginning of the session, the patient is asked to complete the form, and if time permits and if relevant, we might plot the data together to see how the patient is progressing. These measures are meant to be guides. Single-item measures are subject to biases, but the form also helps ensure that my patients and I share similar goals. As one symptom or target behavior improves, or as events occur in the patient's life, we might add or refine our goals. I also include a general assessment of self-care activities that I consider to be important for most patients (exercise, stress management/relaxation/doing something fun, nutrition, social support, sleep hygiene) and treatment use (medication adherence, side effects, use of CBT or other strategies) and, as appropriate, use standard measures.

Target symptom 1: *Nausea*

In the last week, how often did you experience this? (circle an answer)			
never	almost never	a few times a week	a few times a day

In the last week, how severe was this symptom? (circle a number)									
not bad at all			moderate				severe		
1	2	3	4	5	6	7	8	9	10

FIGURE 8–1. Example of a therapist-developed rating to assess therapy goals.

Goal: *To improve my confidence and satisfaction when dating*

In the last week, how satisfied have you been with your relationships? (circle an answer)			
1 = not at all	2 = pretty unsatisfied	3 = pretty satisfied	4 = very satisfied

Goal: *To increase my confidence as a parent*

In the last week, how confident have you felt with your parenting skills? (circle an answer)			
1 = not at all	2 = somewhat	3 = pretty confident	4 = very confident

FIGURE 8–2. Two examples of simple, therapist-developed scales used to measure progress toward therapy goals.

Electronic Office Assessment Systems

In recent years, some software programs have been developed to facilitate measurement-based care and other electronic activities for psychiatric practice. For example, the Psychiatry Quality Measurement (PQM) is a computerized documentation and quality-measurement system (Joubert 2001). According to its developers, the PQM system was designed to document all aspects of a psychiatric patient's treatment, admission, and subsequent follow-up; reduce the time spent on administration; and increase the availability of data from a physician's practice. The system incorporates a program displaying descriptive statistics and can calculate a large range of quality measures, based either on the full population in the database or on preselected subgroups. Data can also be exported for more extensive statistical analysis (Joubert 2001).

One user noted, however, that entering data into the clinical charts is quite time-consuming.

Several other software packages are available for psychiatric practice. The following are examples of programs that have been advertised on the Internet (with no recommendation from me):

- *Valant Psychiatric EMR*—a powerful electronic medical record (EMR) software application that is specialty specific for psychiatrists but can accommodate psychologists as well.
- *A+ DELPHI Psychotherapy Billing Software*—a complete program designed specifically for mental health care professionals. It features extensive note taking for Psychotherapy and Progress Notes and Assessments and Intakes.
- *PatientTrac*—an EMR for psychiatrists and mental health facilities. Designed by board-certified psychiatrists, this EMR does not rely on templates but is a "living" system that can track differential data for subsequent visits.
- *MedEZ*—a medical office management and billing software that has modules designed for mental health centers, psychiatrists, and private practice facilities.
- *Q.D. Clinical EMR*—features hundreds of customizable templates, allowing mental health providers to alter the application to meet their needs.
- *OfficeEMR*—caters to small and mid-sized practices that want a fully integrated EMR and practice management system. OfficeEMR is Web based, Certification Commission for Healthcare Information Technology certified, and Health Insurance Portability and Accountability Act of 1996 compliant.
- *CollaborateMD*—a Web-enabled practice management and medical billing software system with free support and software updates. Requiring only an Internet connection, CollaborateMD manages all data backups and clearinghouse transmissions.

Examples of Electronic Data Systems to Facilitate Evidence-Based Psychiatric Practice

Example 1: A Measurement System to Facilitate Implementation of the STAR*D Guidelines

An example of a potentially useful office-based system is the one developed by Trivedi and Daly (2007) to help implement the STAR*D guidelines. In the STAR*D trial, poorer outcomes were most clearly associated with a failure to change medication after an extended (12–24 weeks) treatment trial that had not produced a meaningful benefit, as well as with a failure to increase the medication dose in a timely manner within the first 12 weeks, despite a lack of significant side effects (Trivedi and Daly 2007). The system developed by Trivedi and Daly (2007) incorporates the most current information about the treatment of depression and provides an easy-to-use interface allowing physicians to use this decision support tool within the context of a routine clinical visit. A computerized algorithm has been developed to facilitate the process of following the suggested dosing schedules and tactical recommendations by displaying the recommended dosages and pretreatment options at that point in time according to the decision. Additionally, all patient information, medication information, medication dosages, next appointments, and progress notes are accessible with a click of the computer mouse and recorded electronically. The program also provides a recommended time frame for the patient to return according to algorithm stage.

Example 2: An Electronic System to Help Monitor Symptom Severity, Plotted Against Clinical Activities

Sentiens has developed an e-health system to provide improved health care for the Western Australian population (see Chapter 30 for an example of how it was used in a clinical case). The system was designed to be used by practitioners to manage and monitor patients wherever they may be—whether close by in Perth, in less populated areas of Western Australia, or working overseas. The resultant HealthSteps system has been implemented in a few research trials, in both Australia (e.g., Barnes et al. 2007) and the United States (e.g., Harvey et al. 2007), to refine further its features and show its efficacy in the management of longer-term health conditions. The Sentiens system can be used to provide real-time data on symptom severity, plotted against clinical activities including medications.

Example 3: Clinical Research Information System at Duke University

Duke University has developed the Clinical Research Information System (CRIS), described as a comprehensive electronic behavioral health care management system (psychiatry.mc.duke.edu/CMRIS/CMRIndex.htm). CRIS was designed to integrate clinical care at all levels. CRIS also employs a clinical rules engine to help guide clinical practices and create a clinical outcomes data warehouse for retrospective decision support. Progress reports can also be generated in real time (http://psychiatry.mc.duke.edu/CMRIS/Image4Symptom.htm).

The cases presented in Parts II and III of this volume also provide examples of how measurement can be used in a variety of clinical settings.

References

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition. Washington, DC, American Psychiatric Association, 1994

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

Anastasia A, Urbina S: Psychological Testing, 7th Edition. New York, Macmillan, 1997

Bagby RM, Ryder AG, Schuller DR, et al: The Hamilton Depression Rating Scale: has the gold standard become a lead weight? Am J Psychiatry 161:2163–2177, 2004

Barnes C, Harvey R, Mitchell P, et al: Evaluation of an online relapse prevention program for bipolar disorder: an overview of the aims and methodology of a randomised controlled trial. Disease Management and Health Outcomes 15:215–224, 2007

Beaton DE, Boers M, Wells GA: Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. Curr Opin Rheumatol 14:109–114, 2002

Beck AT, Ward C, Mendelson M: An inventory for measuring depression. Arch Gen Psychiatry 4:561–571, 1961

Carroll BJ: Why the Hamilton Depression Rating Scale endures. Am J Psychiatry 162:2395–2396, 2005

Cohen J: Statistical Power for the Behavioral Sciences, 2nd Edition. Hillsdale, NJ, Erlbaum, 1988

Corruble E, Hardy P: Why the Hamilton Depression Rating Scale endures. Am J Psychiatry 162:2394, 2005

DeVellis RF: Scale Development: Theory and Applications, 2nd Edition. Thousand Oaks, CA, Sage, 2003

Edwards BC, Lambert MJ, Moran PW, et al: A meta-analytic comparison of the Beck Depression Inventory and the Hamilton Rating Scale for Depression as measures of treatment outcome. Br J Clin Psychol 23 (part 2):93–99, 1984

Ellwood PM: Shattuck lecture outcomes management: a technology of patient experience. N Engl J Med 318:1549–1556, 1988

Fleiss, JL. Statistical Methods for Rates and Proportions. Hoboken, NJ, Wiley Interscience, 1981

Freedland KE, Skala JA, Carney RM, et al: The Depression Interview and Structured Hamilton (DISH): rationale, development, characteristics, and clinical validity. Psychosom Med 64:897–905, 2002

Hamilton M: A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56–62, 1960

Harvey R, Smith M, Abraham N, et al: The Hurricane Choir: remote mental health monitoring of participants in a community based intervention in the post Katrina period. J Health Care Poor Underserved 18:356–361, 2007

Jacobson NS, Truax P: Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 59:12–19, 1991

Jaeschke R, Singer J, Guyatt GH: Measurement of health status: ascertaining the minimal clinically important difference. Control Clin Trials 10:407–415, 1989

Joubert AF: PQM—Psychiatry Quality Measurement. Int Clin Psychopharmacol 16 (suppl 3):S25–33, 2001

Kenardy JA, Dow MGT, Johnston DW, et al: A comparison of delivery methods of cognitive-behavioral therapy for panic disorder: an international multicenter trial. J Consult Clin Psychol 71:1068–1075, 2003

Killen JD, Taylor CB, Hammer LD, et al: An attempt to modify unhealthful eating attitudes and weight regulation practices of young adolescent girls. Int J Eat Disord 13:369–384, 1993

Lambert MJ, Hatch DR, Kingston MD, et al: Zung, Beck, and Hamilton Rating Scales as measures of treatment outcome: a meta-analytic comparison. J Consult Clin Psychol 54:54–59, 1986

Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics 33:159–174, 1977

McGrath BM, Tempier RP: Implementing quality management in psychiatry: from theory to practice—shifting focus from process to outcome. Can J Psychiatry 48:467–474, 2003

Montgomery SA, Åsberg M: A new depression scale designed to be sensitive to change. Br J Psychiatry 134:382–389, 1979

Newman MG, Kenardy J, Herman S, et al: Comparison of cognitive-behavioral treatment of panic disorder with computer assisted brief cognitive-behavioral treatment. J Consult Clin Psychol 65:173–178, 1997

Rush AJ, Giles DE, Schlesser MA, et al: The Inventory for Depressive Symptomatology (IDS): preliminary findings. Psychiatry Res 18:65–87, 1986

Rush AJ, Trivedi MH, Ibrahim HM, et al: The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 54:573–583, 2003

Rush AJ Jr, First MB, Blacker D (eds): Handbook of Psychiatric Measures, 2nd Edition. Washington, DC, American Psychiatric Publishing, 2008

Trivedi MH, Daly EJ: Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. Drug Alcohol Depend 88 (suppl 2):S61–S71, 2007

Trivedi MH, Rush AJ, Gaynes BN, et al: Maximizing the adequacy of medication treatment in controlled trials and clinical practice: STAR(*)D measurement-based care. Neuropsychopharmacology 32:2479–2489, 2007

Williams JB: Standardizing the Hamilton Depression Rating Scale: past, present, and future. Eur Arch Psychiatry Clin Neurosci 251 (suppl 2):II6–II12, 2001

9

Diagnostic Tests

Gregory E. Gray, M.D., Ph.D.

Questions about diagnostic tests are most commonly related to the accuracy of the test, which is the focus of this chapter. For information on topics such as the clinical and economic impact of screening, clinical decision rules, and the differential diagnosis process, readers are referred to recent monographs (Hunink et al. 2001; Knottnerus 2002) and standard textbooks of clinical epidemiology (Fletcher et al. 1996; Sackett et al. 1991).

Evaluating Diagnostic Tests

The accuracy of a diagnostic test is generally assessed in a cross-sectional study, in which patients are evaluated with both the "gold standard" diagnostic procedure and the diagnostic test under evaluation (Knottnerus and Muris 2002; Knottnerus et al. 2002; Newman et al. 2001). Several important issues can affect the validity of such evaluations (Fletcher et al. 1996; Knottnerus and Muris 2002; Knottnerus et al. 2002; Newman et al. 2001; Sackett et al. 1991), but the two major issues are the choice of the gold standard and the choice of subjects.

Choice of Gold Standard

The first issue is related to the choice of the gold standard. In some branches of medicine, the gold standard might be a pathological diagnosis based on a biopsy or an autopsy; in psychiatry, we usually rely on the DSM-IV-TR criteria (American Psychiatric Association [APA] 2000).

Although the DSM-IV-TR diagnostic criteria are often used as the gold standard for evaluating diagnostic tests, they have limitations. First, there are unresolved issues concerning the validity of the criteria themselves (Goldstein and Simpson 2002; Kendell 1989; Kendler 1990).

The second issue relates to the method of eliciting symptoms from patients and of using this information to arrive at a diagnosis. In unstructured clinical interviews, clinicians may ignore or not inquire about certain symptoms and may choose not to follow the DSM criteria in arriving at a diagnosis (Robins 2002). As a result, unstructured interviews may not be reliable. For this reason, a variety of diagnostic instruments have been developed to standardize the process of psychiatric diagnosis, including structured clinical interviews (e.g., the Structured Clinical Interview for DSM-IV [SCID]) and structured interviews that may be administered by either a lay interviewer or a computer (e.g., the Composite International Diagnostic Interview [CIDI] and the Diagnostic Interview Schedule [DIS]) (Kobak et al. 2008). Although these standardized instruments increase the reliability of the diagnoses made, issues of validity remain (Murphy 2002; Narrow et al. 2002; Regier et al. 1998; Robins 2002).

For the assessment of particular symptoms or specific measures of cognitive function (e.g., memory or intelligence), well-established instruments are generally used as the gold standard. Descriptions of many of these are included in the *Handbook of Psychiatric Measures* (Rush et al. 2008)

Choice of Subjects

The other major issue relates to the choice of subjects. We administer diagnostic tests because we have uncertainty about a diagnosis. A cross-sectional

study that includes a spectrum of patients who are similar to those to whom the test is expected to be administered in clinical practice is the most appropriate design. All too often, a test is evaluated in a population composed of a mix of very ill patients and a healthy control group. In such a population, the test performs much better in distinguishing the ill from the healthy than in actual practice. This is referred to as *spectrum bias* (Knottnerus et al. 2002; Newman et al. 2001).

Measures of Test Performance: Dichotomous Results

A variety of terms are used to describe the performance of a diagnostic test (Habbema et al. 2002). As an aid in discussing these terms, Table 9–1 displays the results of a hypothetical comparison of a new diagnostic test with an appropriate gold standard. In this section, we consider test results to be dichotomous (i.e., either positive or negative).

True Positive, True Negative, False Positive, and False Negative

If both the diagnostic test being evaluated and the gold standard yield positive results, the result of the diagnostic test is considered true positive (Table 9–1, cell A). Likewise, if both yield negative results, the result of the diagnostic test is considered true negative (Table 9–1, cell D). If the diagnostic test under evaluation gives a positive result, but the gold standard gives a negative result, the result of the diagnostic test is considered false positive (Table 9–1, cell B). Similarly, if the diagnostic test result is negative, but the gold standard result is positive, the result of the diagnostic test is considered false negative (Table 9–1, cell C).

Sensitivity and Specificity

Sensitivity refers to the proportion of patients with the disease (as assessed by the gold standard) who are detected by the diagnostic test. In Table 9–1, this is calculated as A/(A+C). A highly sensitive test is one that detects most cases of disease.

Specificity refers to the proportion of patients without the disease (as assessed by the gold standard) who are identified by the diagnostic test as *not* having the disease. In Table 9–1, this is calculated as D/(B+D). A highly specific test is one that does not misidentify healthy individuals as having disease.

Although it may seem counterintuitive, highly sensitive diagnostic tests are most useful for ruling *out* diseases. This is because such tests seldom miss cases of disease. Sackett et al. (2000) have proposed the mnemonic SnNout (**Sn**sitive test, **N**egative result, rules **out** disease) as a teaching aid. Similarly, highly specific tests are most useful for ruling *in* diseases, because they seldom misidentify healthy individuals. The mnemonic for such tests is SpPin (**Sp**ecific test, **P**ositive test, rules **in** disease).

Confidence intervals (CIs) for specificity and sensitivity can be calculated using the tables given by Habbema et al. (2002).

Positive and Negative Predictive Values

The positive predictive value (PPV) is the proportion of positive test results that is true positives. Using Table 9–1, PPV is calculated as A/(A+B). The negative predictive value (NPV) is the proportion of negative test results that is true negatives. Using Table 9–1, NPV is calculated as D/(C+D).

PPV and NPV can also be looked at in terms of posttest probabilities of disease (Fletcher et al. 1996; Habbema et al. 2002). PPV represents the probability that an individual with a positive test result has the disease, whereas NPV represents the probability that an individual with a negative test result does not have the disease.

PPV and NPV are dependent on the prevalence of disease in the population being tested, which is also referred to as the *pretest probability of disease* (Fletcher et al. 1996). Mathematically, the relationship is as follows:

$$\text{PPV} = \frac{\text{sensitivity} \times \text{prevalence}}{(\text{sensitivity} \times \text{prevalence}) + [(1 - \text{specificity})(1 - \text{prevalence})]}$$

Because of this dependence on disease prevalence, screening for diseases in low-prevalence populations yields few true-positive test results, regardless of the sensitivity and specificity of the test (Fletcher et al. 1996). Baldessarini et al. (1983) have provided an excellent discussion of the effects of disease prevalence on PPV in psychiatric practice, using as their example the dexamethasone suppression test as a diagnostic test for depression.

TABLE 9–1. Possible results of a diagnostic test

	Disease present	Disease absent	Totals
Test results positive	**A** True positive	**B** False positive	**A+B**
Test results negative	**C** False negative	**D** True negative	**C+D**
Totals	**A+C**	**B+D**	

Sensitivity=A/(A+C)
Specificity=D/(B+D)
Positive predictive value=A/(A+B)
Negative predictive value=D/(C+D)
Likelihood ratio of positive test=[A/(A+C)]/[B/(B+D)]
Likelihood ratio of negative test=[C/(A+C)]/[D/(B+D)]
Error rate=(B+C)/(A+B+C+D)
Accuracy=(A+D)/(A+B+C+D)

Likelihood Ratios

Although PPV and NPV allow the clinician to estimate the probability of disease in a patient with a particular test result, they are both dependent on the prevalence of disease in the study population. Likelihood ratios (LRs), in contrast, are independent of disease prevalence.

The likelihood ratio of a positive test (LR+) is the ratio of the likelihood (probability) of a positive test result in the population with disease divided by the likelihood of a positive test result in the population without disease. This can be calculated as follows, using the data in Table 9–1:

$$\text{LR}+ = \frac{\text{A}/(\text{A}+\text{C})}{\text{B}/\text{B}+\text{D}} = \frac{\text{sensitivity}}{1-\text{specificity}}$$

The likelihood ratio of a negative test result (LR–) is similarly the ratio of the likelihood of a negative test result in the population with disease divided by the likelihood of a negative test result in the population without disease. Once again, using the data in Table 9–1:

$$\text{LR}- = \frac{\text{C}/(\text{A}+\text{C})}{\text{D}/(\text{B}+\text{D})} = \frac{1-\text{sensitivity}}{\text{specificity}}$$

Formulas for calculating CIs for LRs are given by Habbema et al. (2002).

One of the most useful features of LRs is that they can be used to estimate the probability that a given patient has an illness, given a particular test result. The following formula can be used for this estimate:

$$\text{posttest odds} = \text{pretest odds} \times \text{LR}$$

The odds of an event occurring is the ratio of the probability of an event occurring (P) divided by the probability of the event not occurring ($1-P$), or the odds $= P/(1-P)$.

The pretest probability of disease in a patient can be estimated from the prevalence of a disease in your own or similar patient populations, as derived from papers evaluating diagnostic tests, epidemiologic studies, or hospital statistics (Mant 1999; Sackett et al. 2000). For some medical conditions, clinical prediction rules may be used to better refine the estimated pretest probability (McGinn 2002). Sometimes published data are not available and a subjective estimate of pretest probability must be made, although this is subject to numerous biases (Elstein and Schwartz 2002; Hunink et al. 2001). When there is uncertainty about the pretest probability, it is best to estimate an upper and lower limit and to do the calculations with both values to see if it will change your management of the patient.

Once you have a pretest probability of disease, it is converted to the pretest odds, using the formula above. The pretest odds are then multiplied by the appropriate likelihood ratio (LR+ if the test is positive or LR– if the test is negative). The resulting posttest odds can then be converted to a posttest probability, using the following formula:

$$P = \frac{\text{odds}}{1+\text{odds}}$$

For those wishing to go directly from pretest probability to posttest probability, without having to go back and forth between probabilities and odds, nomograms and computer programs are available (Fagan 1975; Glasziou 2001).

Using diagnostic tests in this way to quantify the risk of disease in a patient and to then use these probabilities in clinical decision making has become an important field of study, but it is outside the scope this chapter. For further details, readers are referred to Hunink et al. (2001) and Sackett et al. (1991, 2000).

Error Rate and Accuracy

The error rate of a diagnostic test is the percentage of test results that is either false positives or false negatives (Habbema et al. 2002). Using the data in Table 9–1, error rate can be calculated as follows:

$$\frac{B+C}{A+B+C+D}$$

The term *accuracy* is sometimes used to describe the percentage of test results that is either true positives or true negatives (Fletcher et al. 1996; Greenhalgh 2001). Using the data in Table 9–1, accuracy is calculated as follows:

$$\frac{A+D}{A+B+C+D}$$

Accuracy can also be calculated as 1−error rate.

Diagnostic Odds Ratio

The diagnostic odds ratio (DOR) is sometimes used as a measure of overall test performance (Habbema et al. 2002). Using the data in Table 9–1, DOR is calculated as follows:

$$\frac{A \times D}{B \times C}$$

The DOR is sometimes used in systematic reviews and in meta-analyses because of its mathematical properties; however, it is not a statistic that is easy to interpret clinically (Deeks 2001).

Example of Calculations

Table 9–2 provides data from a study by Watkins et al. (2001) that evaluates the accuracy of a single question ("Do you often feel sad or depressed?") in screening for depression in stroke patients. In the study, answers to the question ("yes" or "no") were compared with the results of the Montgomery-Åsberg Rating Depression Rating Scale (MADRS), with a MADRS score >6 indicating depression. As can be seen from Table 9–2, the single question had a sensitivity of 86% and a specificity of 78%.

We can use LRs to see the impact of a positive or negative response to the question on the probability of a patient being depressed. Depression typically occurs in 25%–40% of patients with stroke and other neurologic conditions (APA 2000). Thus, the pretest odds for such a patient are between 25/75=0.33 and 40/60=0.67. If the patient answers "yes" to the question, the posttest odds are between 0.33×3.87=1.28 and 0.67×3.87=2.59, corresponding to posttest probabilities of 56%–72%. If the patient answers "no," the posttest odds are between 0.33×0.18=0.06 and 0.67×0.18=0.12, corresponding to posttest probabilities of 6%–11%.

Measures of Test Performance: Ordinal or Continuous Results

Although we often think of diagnostic test results as positive or negative, rating scales and many laboratory tests have more than just two values. Rating scales, for example, can have outcomes that are ordered categories (e.g., none, minimal, mild, moderate, or severe) or that yield a numerical score. Although a cutoff value is often chosen to convert such an outcome to a dichotomous measure, information is lost in the process.

Choosing a Cutoff Score

When a test result can take more than two values as described above, a decision must be made about a cutoff score if the results are to be viewed as either positive or negative. The choice of the cutoff value involves a trade-off between sensitivity and specificity (Fletcher et al. 1996).

Table 9–3 shows the results of a study of the CAGE questionnaire as a screening tool for diagnosing alcohol abuse in general medical patients (Buchsbaum et al. 1991). (CAGE is an acronym for certain key symptoms of alcohol abuse.) As can be seen, as the cutoff score is increased, the specificity increases, but the sensitivity decreases. At the usual cutoff score of 2 (Martino et al. 2008), the sensitivity is 74% and the specificity is 91%.

Receiver Operator Characteristic Curves

The data in Table 9–3 can also be displayed graphically, in the form of a receiver operator characteristic

TABLE 9–2. Accuracy of a single question ("Do you often feel sad or depressed?") in screening for depression in stroke patients

Answer to question	Depressed (MADRS>6)	Not depressed (MADRS ≤6)	Totals
Yes	37	8	45
No	6	28	34
Totals	43	36	79

Sensitivity=37/43=86%
Specificity=28/36=78%
Positive predictive value=37/45=82%
Negative predictive value=28/34=82%
Likelihood ratio of positive test=0.86/0.22=3.9
Likelihood ratio of negative test=0.14/0.78=0.18
Error rate=14/79=18%
Accuracy=65/79=82%

Note. MADRS=Montgomery-Åsberg Depression Rating Scale.
Source. Data from Watkins et al. 2001.

(ROC) curve (Figure 9–1). An ROC curve plots the sensitivity against (1–specificity) for each cutoff value (Fletcher et al. 1996; Sackett et al. 1991, 2000). The ROC curve can be used to identify a cutoff value, where a point near the upper left corner is the appropriate cutoff. ROC curves can also be used to compare two diagnostic tests; the better test is the one with its ROC curve closer to the upper left corner (Fletcher et al. 1996).

Likelihood Ratios

Intuitively, it would seem that converting numerical or ordered test results into just two categories would result in a loss of information. It would seem that a value greatly over the cutoff would be more indicative of disease than a value that is barely more than the cutoff. LRs allow these differences to be taken into account.

If a test result takes multiple values, LRs can be calculated for multiple scores, not just LR+ and LR–. This is illustrated in Table 9–4, using the same CAGE data found in Table 9–3 (Buchsbaum et al. 1991). It can be seen that as the CAGE score increases from 0 to 4, LRs increase from 0.14 to 100. Assuming a prevalence (pretest probability) of alcohol dependence of 5% (APA 2000), a CAGE score of 0 decreases the posttest probability to 1%. Using the same figures, a CAGE score of 2 increases the posttest probability to 19% and a CAGE score of 5 increases the posttest probability to 84%. Clearly, this provides more information than simply treating the results as *positive* (2+) or *negative* (0–1).

Systematic Reviews and Meta-Analyses of Diagnostic Tests

As with the evaluation of therapies, systematic reviews that are related to diagnostic tests may be conducted. However, there are several approaches to meta-analysis (Deeks 2001; Glasziou et al. 2001). One approach is to use one forest plot for sensitivities and their CIs and another for specificities and their CIs (see Chapter 6). Another approach is to plot the sensitivity against (1–specificity) on an

TABLE 9–3. Sensitivity and specificity of the CAGE questionnaire in screening for alcohol abuse in general medical patients

CAGE score cutoff	Sensitivity (%)	Specificity (%)
0 vs. 1+	89	81
<2 vs. 2+	74	91
<3 vs. 3+	44	98
<4 vs. 4+	25	100

Source. Data from Buchsbaum et al. 1991.

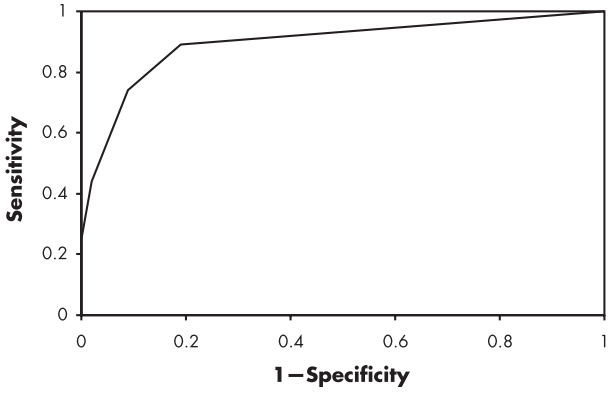


FIGURE 9–1. Receiver operator characteristic curve for CAGE questionnaire.

Source. Data from Buchsbaum et al. (1991).

ROC plot, with each study representing one point on the plot. Methods deriving pooled estimates of sensitivity, specificity, and LRs are also available. Finally, DORs can be pooled and this can be used to derive a summary ROC curve. Although such methods can demonstrate that a test is useful, if there is heterogeneity in results the summary statistics may not give a valid estimate of the probability associated with a specific test outcome (Deeks 2001).

Critical Appraisal Guide for Studies of Diagnostic Test Accuracy

Having discussed how diagnostic tests are evaluated and measures of their accuracy, we now turn to the critical appraisal of reports of such evaluations. A guide for critically appraising studies of diagnostic test accuracy is given in Table 9–5.

Is the Study Valid?

As discussed earlier in this chapter, the appropriate study design to assess the accuracy of a diagnostic test is a cross-sectional study in which patients are evaluated with both the gold standard diagnostic procedure and the diagnostic test under evaluation (Knottnerus and Muris 2002; Knottnerus et al. 2002; Newman et al. 2001). Both tests should be administered blindly to avoid observer bias (Fletcher et al. 1996), and the results of one test should not influence the decision to administer the other test. The gold standard itself should be appropriate, and an appropriate spectrum of patients should be chosen to avoid spectrum bias (Knottnerus et al. 2002; Newman et al. 2001).

TABLE 9–4. Likelihood ratios for specific CAGE scores in screening for alcohol abuse in general medical patients

CAGE score	Likelihood ratio
0	0.14
1	1.5
2	4.5
3	13
4	100

Source. Data from Buchsbaum et al. 1991.

Are the Results Important?

Sensitivity and specificity are the usual measures of diagnostic test accuracy. Which of the two is more important will depend on the purpose of the test. Recall that highly sensitive diagnostic tests are most useful for ruling *out* diseases (SnNout: **Sen**sitive test, **N**egative result, rules **out** disease), whereas highly specific tests are most useful for ruling *in* diseases (SpPin: **Sp**ecific test, **P**ositive test, rules **in** disease) (Sackett et al. 2000).

PPV and NPV will provide estimates of the probabilities of disease in patients with positive or negative test results, respectively, in the study population. However, LR+ and LR– are more useful measures, because they can be generalized to other populations with different disease prevalences (Fletcher et al. 1996; Mant 1999).

Can I Apply the Results to My Patient?

The first question to ask is whether the test is feasible in your setting. Does it require specialized equipment or is it very expensive? Does it require expertise to administer or to interpret the test, and is such expertise available?

The next question is specific to your patient and relates to whether the results of the test will change your patient management. As described in more detail elsewhere (Hunink et al. 2001; Sackett et al. 1991, 2000), physicians make decisions about clinical management, based on their (often unstated) as-

TABLE 9–5. Critical appraisal guide for studies of diagnostic accuracy

Is the study valid?
Is it a cross-sectional study?
Was the test evaluated in an appropriate spectrum of patients?
Was there an independent blind comparison with a "gold standard"?
Were both tests administered, regardless of outcome?
Are the results important?
What is the sensitivity?
What is the specificity?
What is the positive predictive value?
What is the negative predictive value?
What is the likelihood ratio for a positive test result?
What is the likelihood ratio for a negative test result?
Can I apply the results to my patient?
Is the test available or feasible in my setting?
Do I have enough information to apply the test and to interpret the results?
Given reasonable pretest estimates of disease probability in my patient, what are the posttest probabilities if the test is positive? If it is negative?
Will these probabilities change my management of the patient?

Source. Based on Gilbert et al. 2001; Glasziou et al. 2001; Sackett et al. 2000.

sessments of the probability of disease in a given patient. If that probability is high, physicians treat the patient; if that probability is intermediate, physicians order more diagnostic tests; and if it is low, physicians do neither. The probability at or above which tests are ordered is termed the *test threshold*, and the probability at or above which treatment is begun is termed the *treatment threshold* (Hunink et al. 2001; Sackett et al. 2000).

If your patient's pretest probability is already above the treatment threshold, there is no point in ordering the diagnostic test. For example, if your diagnostic interview shows that your patient met the DSM-IV-TR criteria for major depression, there is no need to also use a rating scale developed to screen for depression. You have already established the diagnosis with enough certainty to begin treatment.

Conversely, if the posttest probability remains low even if the test result is positive, the test should probably not be performed because it is likely that any positive test result will be a false positive. Ordering the test under these circumstances would be both a waste of money and an unnecessary cause of anxiety for patients with false-positive results (Fletcher et al. 1996).

To do these sorts of calculations, you will need the LR+ and LR– from the article reporting the performance of the test in question and an estimate of your patient's pretest probability of disease. If your patient is similar to patients in the study population, the pretest probability can be derived from the study evaluating the diagnostic test. If not, the pretest probability can be estimated from epidemiologic data or from patient data from your hospital or health maintenance organization (Mant 1999; Sackett et al. 2000). If published data are not available, you may have to make a subjective estimate of the pretest. In this case, it is probably best to estimate an upper and lower limit and to do the calculations with both values.

If the test is valid and feasible and if the results will make a difference in patient management, you should use it.

References

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

Baldessarini RJ, Finkelstein S, Arana GW: The predictive power of diagnostic tests and the effect of prevalence of illness. Arch Gen Psychiatry 40:569–573, 1983

Buchsbaum DG, Buchanan RG, Centor RM, et al: Screening for alcohol abuse using CAGE scores and likelihood ratios. Ann Intern Med 115:774–777, 1991

Deeks JJ: Systematic reviews of evaluation of diagnostic and screening tests. BMJ 323:157–162, 2001

Elstein AS, Schwartz A: Clinical problem solving and diagnostic decision making: a selective review of the cognitive research literature, in The Evidence Base of Clinical Diagnosis. Edited by Knottnerus JA. London, BMJ Books, 2002, pp 179–195

Fagan TJ: Nomogram for Bayes' theorem. N Engl J Med 293:257, 1975

Fletcher RH, Fletcher SW, Wagner EH: Clinical Epidemiology: The Essentials, 3rd Edition. Baltimore, MD, Williams & Wilkins, 1996

Gilbert R, Logan S, Moyer VA, et al: Assessing diagnostic and screening tests, II: how to use the research literature on diagnosis. West J Med 175:37–41, 2001

Glasziou P: Which methods for bedside Bayes? ACP J Club 135:A11–A12, 2001

Glasziou P, Irwig L, Bain C, et al: Systematic Reviews in Health Care: A Practical Guide. Cambridge, UK, Cambridge University Press, 2001

Goldstein JM, Simpson JC: Validity: definitions and applications to psychiatric research, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 149–163

Greenhalgh T: How to Read a Paper: The Basics of Evidence Based Medicine, 2nd Edition. London, BMJ Books, 2001

Habbema JDF, Eijkemans R, Krijnen P, et al: Analysis of data on the accuracy of diagnostic tests, in The Evidence Base of Clinical Diagnosis. Edited by Knottnerus JA. London, BMJ Books, 2002, pp 117–143

Hunink M, Glasziou P, Siegel J, et al: Decision Making in Health and Medicine: Integrating Evidence and Values. New York, Cambridge University Press, 2001

Kendell RE: Clinical validity. Psychol Med 19:45–55, 1989

Kendler KS: Toward a scientific psychiatric nosology: strengths and limitations. Arch Gen Psychiatry 47:969–973, 1990

Knottnerus JA (ed): The Evidence Base of Clinical Diagnosis. London, BMJ Books, 2002

Knottnerus JA, Muris JWM: Assessment of the accuracy of diagnostic tests: the cross-sectional study, in The Evidence Base of Clinical Diagnosis. Edited by Knottnerus JA. London, BMJ Books, 2002, pp 39–60

Knottnerus JA, van Weel C, Muris JWM: Evaluation of diagnostic procedures. BMJ 324:477–480, 2002

Kobak KA, Skodol AE, Bender DS: Diagnostic measures for adults, in Handbook of Psychiatric Measures, 2nd Edition. Edited by Rush AJ, First MB, Blacker D. Washington, DC, American Psychiatric Publishing, 2008, pp 35–60

Mant J: Is this test effective? In Evidence-Based Practice: A Primer for Health Professionals. Edited by Dawes M, Davies P, Gray A, et al. New York, Churchill Livingstone, 1999, pp 133–157

Martino S, Poling J, Rounsaville BJ: Substance use disorders measures, in Handbook of Psychiatric Measures, 2nd Edition. Edited by Rush AJ, First MB, Blacker D. Washington, DC, American Psychiatric Publishing, 2008, pp 437–470

McGinn T: Practice corner: using clinical prediction rules. ACP J Club 137:A10–A11, 2002

Murphy JM: Symptom scales and diagnostic schedules in adult psychiatry, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 273–332

Narrow WE, Rae DS, Robins LN, et al: Revised prevalence estimates of mental disorders in the United States: using a clinical significance criterion to reconcile 2 surveys' estimates. Arch Gen Psychiatry 59:115–123, 2002

Newman TB, Browner WS, Cummings SR: Designing studies of medical tests, in Designing Clinical Research: An Epidemiologic Approach, 2nd Edition. Edited by Hulley SB, Cummings SR, Browner WS, et al. Philadelphia, PA, Lippincott Williams & Wilkins, 2001, pp 175–191

Regier DA, Kaelber CT, Rae DS, et al: Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. Arch Gen Psychiatry 55:109–115, 1998

Robins LN: Birth and development of psychiatric interviews, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 257–271

Rush AJ, First MB, Blacker D (eds): Handbook of Psychiatric Measures, 2nd Edition. Washington, DC, American Psychiatric Publishing, 2008

Sackett DL, Haynes RB, Guyatt GH, et al: Clinical Epidemiology: A Basic Science for Clinical Medicine, 2nd Edition. Boston, MA, Little, Brown, 1991

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Watkins C, Daniels L, Jack C, et al: Accuracy of a single question in screening for depression in a cohort of patients after stroke: comparative study. BMJ 323:1159, 2001

10

Surveys of Disease Frequency

Gregory E. Gray, M.D., Ph.D.

Questions about disease frequency take various forms. A patient's family member or member of the public may wonder about the risk of developing a disease. An administrator may wonder about the number of patients who will need treatment. A clinician attempting to evaluate a patient's risk of disease, based on the results of a diagnostic test, will need information on the patient's pretest probability of disease (see Chapter 9).

All of these questions require quantitative information about the frequency of a particular illness in a defined population of interest, which is the subject of descriptive epidemiology. This chapter reviews the various measures of disease frequency and the methods of descriptive epidemiology, as background information for the critical appraisal of studies of disease frequency. For additional information, readers are referred to standard epidemiology texts (Fletcher et al. 1996; Gordis 2000; Hennekens et al. 1987; Lilienfeld and Stolley 1994) and reviews of psychiatric epidemiology (Regier and Burke 2000; Tsuang and Tohen 2002).

Measures of Disease Frequency

A number of terms, such as *incidence*, *prevalence*, and *lifetime prevalence*, are used—and often misused—to refer to the frequency of disease in a population.

Incidence

The *incidence* of disease is the number of new cases of disease in a defined population in a given period of time. It is measured by starting with a population (cohort) that is initially free of the disease in question and by then counting the cases that develop in a defined period of time. The incidence rate has "cases" as its numerator and "person-years at risk " as its denominator. The incidence rate is the measure of frequency of most interest to epidemiologists seeking to understand the cause of a disease, because it measures the actual risk of developing the disease.

Point Prevalence

When epidemiologists use the term *prevalence*, they are usually referring to the *point prevalence* (i.e., the number of cases of disease in a defined population at a given point in time). It is measured in a cross-sectional survey in which individuals are identified as either having or not having a disease at the time of the survey.

Point prevalence is of interest to health care planners, because it measures the burden of illness in a population. It is also of interest to clinicians, because it is a measure of a patient's pretest probability of disease (see Chapter 9).

Point prevalence is less helpful in understanding the causes of disease, because it reflects both the incidence (risk) of disease and the duration of the illness. In a *steady state* situation in which the incidence rate and average duration of illness are constant, the following relationship between incidence and point prevalence exists:

$$\text{point prevalence} = \text{incidence rate} \times \text{average duration of illness}$$

This relationship allows for the calculation of the third variable in the equation, if the other two variables are known.

Period Prevalence

Period prevalence is a measure of the proportion of individuals in a defined population who have an illness in a specified period of time. It is a hybrid measure that reflects both the proportion of patients who have the illness at the start of the time interval (prevalence), as well as the number of new cases that develop during the time period (incidence × time).

Although period prevalence is seldom used outside psychiatry (Hennekens et al. 1987), it is widely used in surveys of the frequency of psychiatric illnesses (Fleming and Hsieh 2002; Regier and Burke 2000). In the National Comorbidity Survey, for example, results were reported as 12-month prevalences, which indicates that an individual had the disorder in question at some time during the 12 months prior to the interview date (Kessler and Walters 2002).

The popularity of period prevalence in psychiatric epidemiology relates to the fact that many psychiatric illnesses have a recurring or episodic course; therefore, the period prevalence is believed to be a better reflection of disease burden in the population than is the proportion of patients who are symptomatic at a given point in time.

Lifetime Prevalence

Lifetime prevalence is a specific type of period prevalence used in psychiatric epidemiology. As used in the National Comorbidity Survey and in similar studies, *lifetime prevalence* is the proportion of the patients in population studies who have experienced an illness up to the point in their life that they are surveyed.

Defined in this way, lifetime prevalence has several limitations (Fleming and Hsieh 2002; Regier and Burke 2000). First, it is dependent on memory (recall bias). Second, it is highly dependent on the age structure of the population, with young people less likely to develop the disease than older people. Third, if the disease is associated with excess mortality from suicide or other causes, lifetime prevalence may decline with age.

Morbidity Risk

An alternative measure of lifetime risk, sometimes referred to as *lifetime expectancy*, is morbidity risk or morbid risk (Faraone et al. 2002; Slater and Cowie 1971). This measure corrects the denominator to reflect the fact that younger individuals have had less chance to develop the illness than older individuals and is a better estimate of the risk of developing the illness at some point during a normal lifespan.

Age- and Sex-Specific and Adjusted Rates

The incidence and prevalence of many psychiatric disorders vary with age and sex. The incidence of major depression, for example, is greater in women than in men (Horwath et al. 2002), and the incidence and prevalence of dementia increase with age, beginning at about age 65 years (Hybels and Blazer 2002). In determining the pretest probability of disease for an individual, using age- and sex-specific prevalences will be more accurate than using the prevalence figures for the population as a whole.

When comparing rates between geographic areas or over time, it is important to take demographic differences into account. For example, differences in the incidence of depression in two populations could reflect actual differences in risk for disease or they could simply reflect differences in age or sex distribution. To eliminate demographic differences as the explanation, rates can be standardized. In essence, this involves calculating a weighted average of the age- and sex-specific rates, using identical weights for both populations. Methods of standardization and details of the calculations can be found in standard epidemiology and biostatistics texts (Gordis 2000; Hennekens et al. 1987; Pagano and Gauvreau 2000).

Using Registries and Existing Records

There are several potential sources of existing or routinely collected information about the frequency of disease in a population. These resources include disease registries, reporting systems, and insurance or health plan statistics (Gordis 2000; Lilienfeld and Stolley 1994).

For some diseases (e.g., cancer and some infectious diseases), established disease reporting systems and case registries can provide considerable information about incidence or prevalence. In psychiatry, largely because of privacy concerns, such registers or reporting systems are largely nonexistent. Notable exceptions to this include case registers of psychiat-

ric cases in Rochester, New York; in Denmark; and in Mannheim, Germany (Regier and Burke 2000). Information from such registers is largely limited to schizophrenia and to other serious mental illnesses that require hospitalization.

Records maintained by a health maintenance organization or a health plan can be used to provide useful information about the treated prevalence of mental illness in the health plan's members. Unfortunately, much illness goes untreated, so such records underestimate the actual frequency of mental illness in the health plan's members. As a result, surveys are required to better estimate the actual frequency of disease in a population.

Surveys of Disease Frequency

Two aspects of the design of surveys that are used to determine the incidence or prevalence of a disease in a population are of particular importance: who to survey and how to assess the presence of absence of disease.

Populations and Samples

The first decision involves deciding on the population of interest (Hulley et al. 2001). Although we often are interested in the frequency of disease in the general population, the prevalence of many psychiatric disorders is greater in specific settings (e.g., primary care settings or jails) and can vary with ethnicity (Burke 2002). Thus, there is a rationale for conducting surveys among subpopulations and in specific settings, as well as in the general population.

After deciding on the population of interest, the next step is to assemble a sample of individuals to actually study. A variety of techniques can be used for this purpose (Hulley et al. 2001). If the population is such that all members can be enumerated (e.g., members of a health plan), a simple random sample can be used. For community surveys, cluster sampling is often used. In cluster sampling, the investigators first randomly select census tracts or similar geographic areas, followed by a sample of addresses within the geographic area. Details of the various sampling methods, including methods of determining the sample size required for a certain precision in the estimate of the prevalence rate, can be found in standard texts on survey methods (Kalton 1983; Kelsey et al. 1996).

Assessment Instrument

The second decision relates to the method used to decide whether an individual has the disease of interest or not. Although we use the DSM-IV-TR diagnostic criteria (American Psychiatric Association [APA] 2000), we must also have a method of eliciting symptoms and for applying these criteria during a survey. The chosen method must be feasible, reliable, and valid (Regier and Burke 2000).

Reliability refers to the ability of an assessment instrument to yield a consistent result. In unstructured clinical interviews, clinicians may ignore or not inquire about certain symptoms and may choose not to follow DSM criteria in arriving at a diagnosis (Robins 2002). As a result, unstructured interviews may not be reliable. As noted in Chapter 9, a variety of diagnostic instruments have been developed to standardize and improve the reliability of the psychiatric diagnosis process, including structured interviews, administered either by a clinician (e.g., the Structured Clinical Interview for DSM-IV [SCID]) or by a lay interviewer (e.g., the Composite International Diagnostic Interview [CIDI] and the Diagnostic Interview Schedule [DIS]) (Kobak et al. 2008).

The other issue is *validity* (i.e., whether the instrument measures its intended parameters). Although the standardized instruments mentioned above have improved the reliability of the diagnoses made in surveys, issues of validity remain (Murphy 2002; Regier et al. 1998; Robins 2002). In particular, community surveys identify individuals who have less severe symptoms than those seen among individuals in treatment settings, and it is unclear if these represent milder cases of the same disorders or simply nonpathological transient responses to stressors (Narrow et al. 2002; Regier et al. 1998).

For the assessment of particular symptoms or specific measures of cognitive function (e.g., memory or intelligence), a variety of well-established instruments that are considered both reliable and valid are available. Descriptions of many of these are included in the *Handbook of Psychiatric Measures* (Rush et al. 2008).

Critical Appraisal of a Survey of Disease Frequency

Having discussed the various measures of disease frequency and how they are determined, we now turn to the critical appraisal of reports of disease

frequency. A guide for critically appraising these studies is given in Table 10–1.

Is the Study Valid?

The first questions relate to the study design. For estimating prevalence, a cross-sectional survey, in which individuals are identified as either having or not having a disease either at the time of the survey or during some defined time period, is the appropriate design. Incidence rates are measured in cohort studies, in which a group of individuals who are initially free of the disease in question is followed over time, and cases of disease that develop in the cohort during the follow-up period are counted.

Regardless of study design, the individuals surveyed should be members of a defined population that is sampled in such a way as to give an unbiased estimate or incidence or prevalence rate. A convenience sample or the use of volunteers recruited from an advertisement cannot be considered representative of a defined population; it will give a biased estimate of disease frequency, which is referred to as *selection bias* (Daly and Bourke 2000).

Nonresponse bias is an additional concern, because individuals who refuse to participate in the survey may be systematically different from individuals who do participate (Aday 1996; Barker et al. 1998; Daly and Bourke 2000). For example, telephone follow-up of nonresponders in the National Comorbidity Survey found that they were more apt to have an anxiety disorder than were responders (Kessler and Walters 2002). In general, it is considered desirable to have as high a response rate as possible—generally, at least 75%–80% (Aday 1996; Barker et al. 1998; Kelsey et al. 1996).

It is essential to use standard criteria to decide whether an individual has a disorder or not, because prevalence will depend on the definition of what constitutes a case (Fletcher et al. 1996). In addition, as described above under "Surveys of Disease Frequency," the assessment instrument must have acceptable reliability and validity (Regier and Burke 2000).

Are the Results Important?

Which results are important will, of course, be related to your initial clinical question. At a minimum, you would expect to see an incidence or prevalence rate for each disorder surveyed, with confidence limits. If your goal is to apply the results to an individual patient, rates specific to age, sex, and other demographic characteristics will be helpful. In contrast, if your interest lies in comparing the rates in two populations, standardized rates will be of more interest.

TABLE 10–1. Critical appraisal guide for surveys of disease frequency

Is the study valid?
Is the study design appropriate (cross-sectional survey for measuring prevalence or cohort study for measuring incidence)?
Was an appropriate sampling method used?
Was the response rate sufficiently high?
Was a standardized method used to determine the presence of disease?
Is the assessment instrument reliable and valid?
Are the results important?
What is the incidence or prevalence?
Are there important differences by age, sex, ethnicity, etc.?
How precise are the estimates?
Can I apply the results to my patient?
Does my patient come from a population similar to that surveyed?
Are there age- or sex-specific estimates (if appropriate) that apply to my patient?

Source. Based in part on the criteria of Barker et al. 1998; Fletcher et al. 1996.

Can I Apply the Results to My Patient?

The first question to ask is whether the population is similar enough that the results can reasonably be generalized to your patient (Hulley et al. 2001). In addition, if your clinical question relates to a pretest probability of disease in your patient (see Chapter 9), the availability of age- and sex-specific prevalence rates will give you a more accurate estimate than will crude or standardized rates.

References

Aday LA: Designing and Conducting Health Surveys: A Comprehensive Guide, 2nd Edition. San Francisco, CA, Jossey-Bass, 1996

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

Barker DJP, Cooper C, Rose G: Epidemiology in Medical Practice, 5th Edition. New York, Churchill Livingstone, 1998

Burke JD: Mental health services research, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 165–179

Daly LE, Bourke GJ: Interpretation and Uses of Medical Statistics, 5th Edition. Oxford, UK, Blackwell Scientific, 2000

Faraone SV, Tsuang D, Tsuang MT: Methods in psychiatric genetics, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 65–130

Fleming JA, Hsieh C-C: Introduction to epidemiologic research methods, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 3–33

Fletcher RH, Fletcher SW, Wagner EH: Clinical Epidemiology: The Essentials, 3rd Edition. Baltimore, MD, Williams & Wilkins, 1996

Gordis L: Epidemiology, 2nd Edition. Philadelphia, PA, WB Saunders, 2000

Hennekens CH, Buring JE, Mayrent SL: Epidemiology in Medicine. Boston, MA, Little, Brown, 1987

Horwath E, Cohen RS, Weissman MM: Epidemiology of depressive and anxiety disorders, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 389–426

Hulley SB, Newman TB, Cummings SR: Choosing the study subjects: specification, sampling, and recruitment, in Designing Clinical Research: An Epidemiologic Approach, 2nd Edition. Edited by Hulley SB, Cummings SR, Browner WS, et al. Philadelphia, PA, Lippincott Williams & Williams, 2001, pp 25–36

Hybels CF, Blazer DG: Epidemiology and geriatric psychiatry, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 603–628

Kalton G: Introduction to Survey Sampling. Newbury Park, CA, Sage, 1983

Kelsey JL, Whittemore AS, Evans AS, et al (eds): Methods in Observational Epidemiology, 2nd Edition. New York, Oxford University Press, 1996

Kessler RC, Walters E: The National Comorbidity Survey, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 343–362

Kobak KA, Skodol AE, Bender DS: Diagnostic measures for adults, in Handbook of Psychiatric Measures, 2nd Edition. Edited by Rush AJ, First MB, Blacker D. Washington, DC, American Psychiatric Publishing, 2008, pp 35–60

Lilienfeld DE, Stolley PD: Foundations of Epidemiology, 3rd Edition. New York, Oxford University Press, 1994

Murphy JM: Symptom scales and diagnostic schedules in adult psychiatry, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 273–332

Narrow WE, Rae DS, Robins LN, et al: Revised prevalence estimates of mental disorders in the United States: using a clinical significance criterion to reconcile 2 surveys' estimates. Arch Gen Psychiatry 59:115–123, 2002

Pagano M, Gauvreau K: Principles of Biostatistics, 2nd Edition. Pacific Grove, CA, Duxbury, 2000

Regier DA, Burke JD: Epidemiology, in Kaplan and Sadock's Comprehensive Textbook of Psychiatry, 7th Edition. Edited by Sadock BJ, Sadock VA. Philadelphia, PA, Lippincott Williams & Wilkins, 2000, pp 500–522

Regier DA, Kaelber CT, Rae DS, et al: Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. Arch Gen Psychiatry 55:109–115, 1998

Robins LN: Birth and development of psychiatric interviews, in Textbook in Psychiatric Epidemiology, 2nd Edition. Edited by Tsuang MT, Tohen M. New York, Wiley-Liss, 2002, pp 257–271

Rush AJ, First MB, Blacker D (eds): Handbook of Psychiatric Measures, 2nd Edition. Washington, DC, American Psychiatric Publishing, 2008

Slater E, Cowie V: The Genetics of Mental Disorder. New York, Oxford University Press, 1971

Tsuang MT, Tohen M (eds): Textbook in Psychiatric Epidemiology, 2nd Edition. New York, Wiley-Liss, 2002

This page intentionally left blank

11

Studies of Risk or Harm

Gregory E. Gray, M.D., Ph.D.

This chapter covers the appraisal of two types of studies: those that assess the *harm* associated with a therapy (e.g., medication side effects) and those that attempt to identify factors that increase the *risk* of developing a disease. A variety of study designs are presented, including case reports, randomized controlled trials (RCTs), cohort studies, and case-control studies (Levine et al. 2002; Turcotte et al. 2001) (Table 11–1).

Case Reports and Case Series

As noted in Chapter 5, case reports and case series are best viewed as sources of hypotheses for further testing (Fletcher et al. 1996; Hennekens et al. 1987). They have been used traditionally to describe rare clinical events and often provide the first warnings about rare—but clinically significant—adverse drug reactions (ADRs) (Brewer and Colditz 1999).

Unfortunately, case reports and case series may also include coincidental occurrences, and without a denominator it is impossible to determine the actual risk of an adverse event. For example, case reports have wrongly implicated beta-blockers as a cause of depression, which actually occurs no more frequently in patients on beta-blockers than in patients receiving placebos (Ko et al. 2002).

Randomized Controlled Trials

Although it would be unethical to conduct an RCT designed to cause harm to a subject, a considerable amount of information about common side effects of medications comes from Phase II and Phase III clinical trials (Jones 2001). Unfortunately, such reporting is often incomplete (Ioannidis and Lau 2001). Furthermore, premarketing trials are often of short duration, are limited to fewer than 2,000 patients, and simply do not have the statistical power to detect ADRs that occur in 1 of 10,000 drug exposures (for which 30,000 patients would have to be studied) (Brewer and Colditz 1999; Pirmohamed et al. 1998).

For common side effects, however, an RCT provides the best evidence that the side effect was caused by medication, not chance (Turcotte et al. 2001). A variety of statistical measures are used to assess the magnitude of the difference in adverse events between a control and experimental treatments. Most of these measures are discussed in Chapter 5; however, there are three measures that are unique to studies of harm: relative risk increase (RRI), absolute risk increase (ARI), and number needed to harm (NNH) (Sackett et al. 2000) (Table 11–2).

Relative Risk and Relative Risk Increase

If a therapy increases the risk of an adverse event, the experimental event rate (EER) is greater than the control event rate (CER), and relative risk (RR), calculated as EER/CER, is greater than 1. RRI is calculated using the following formula:

$$\text{RRI} = \frac{\text{EER} - \text{CER}}{\text{CER}} = \text{RR} - 1$$

Because RR and RRI are relative measures, they do not provide an estimate of the actual increase in frequency of an adverse event. For example, an EER

TABLE 11–1. Study designs used to assess etiology or harm

	Case reports or case series	Randomized controlled trial (RCT)	Cohort study	Case-control study
Advantages	Used to describe rare or unusual events	Least biased study design Can demonstrate cause-effect relationship	Less apt to be biased than case-control study Better able to assess rare outcomes than RCT	Quicker, less expensive than cohort study Useful for rare diseases
Disadvantages	Cannot prove association or causality Often misleading	Cannot assess risk of rare side effects Unethical to conduct trial specifically to cause harm	Involves large numbers of subjects Loss to follow-up may limit validity	More prone to bias than RCT or cohort study Cannot measure absolute risks
Risk statistics	None	Relative risk increase Absolute risk increase Number needed to harm	Relative risk Risk difference Number needed to harm	Odds ratio

of 0.5% and a CER of 0.1% yield the same RR and RRI as an EER of 50% and an CER of 10%, yet the absolute magnitude of the increase in adverse events is greater in the latter case.

Absolute Risk Increase and Number Needed to Harm

ARI is calculated as EER – CER. Unlike RR and RRI, ARI provides a measure of the proportion of patients receiving treatment who will be harmed by it, which thus avoids some of the limitations of the relative measures of effect.

NNH is simply the reciprocal of ARI. It is analogous to NNT (see Chapter 5) as a measure of treatment effect; it indicates the number of patients who would need to be treated with the experimental treatment to produce one more adverse event than would have occurred with the control treatment.

Epidemiologic Studies

Because we do not ordinarily conduct experiments to prove that something is harmful to patients, much of our knowledge about the etiology of disease, as well as of less common medication side effects, comes from epidemiologic studies. In this section, the two most common study designs, cohort studies and case-control studies, are reviewed. The reader is referred to several excellent epidemiology texts for a more complete discussion (Fletcher et al. 1996; Gordis 2000; Hennekens et al. 1987; Kelsey et al. 1996; Rothman 2002).

Cohort Studies

In a cohort study, a group of subjects (cohort) is followed over time, and the incidence of disease is determined (see Chapter 10). The members of the cohort can be classified as either *exposed* or *unexposed* to a medication or suspected risk factor for disease. Such characteristics are determined at the beginning of the study, before disease occurrence. If exposure increases the risk of disease, the incidence in the exposed group is greater than the incidence in the unexposed group.

TABLE 11–2. Measures of risk: randomized controlled trial

	Experimental treatment	Control treatment
Harm	A	B
No harm	C	D
Total	A+C	B+D

Experimental event rate (EER) = A/(A+C)
Control event rate (CER) = B/(B+D)
Relative risk (RR) = EER/CER
Relative risk increase = (EER – CER)/CER = RR – 1
Absolute risk increase (ARI) = EER – CER
Number needed to harm = 1/ARI

Prospective Cohort Studies

Cohort studies can be either prospective or retrospective (Hennekens et al. 1987). In a prospective or concurrent cohort study, the cohort is assembled and the exposure status is determined; it is then followed over time to determine the incidence of disease. In a prospective cohort study, Takeshita et al. (2002) followed a cohort of more than 3,000 Japanese Americans for 6 years and found that those with depressive symptoms had an increased mortality rate, compared with those without depressive symptoms.

Retrospective Cohort Studies

In a retrospective cohort study, existing records are used to identify a historical cohort and to measure both exposures and outcomes of interest, all of which have occurred at the time the study is initiated. Doing so allows the study to be conducted more quickly and with less expense than if the cohort was actually followed for several years. Because retrospective cohort studies are dependent on existing records, issues of data quality often arise. An example of a retrospective cohort study is that of Gunnell et al. (2002), in which intellectual performance at age 18 was determined from military psychological records for a cohort of nearly 200,000 Swedish male conscripts examined several years earlier and was found to predict the subsequent development of schizophrenia and other psychoses over an average follow-up period of 5 years.

Bias and Confounding

Cohort studies are sometimes referred to as *natural experiments*, although they differ significantly in that the subjects are not randomly allocated in a cohort study (Gordis 2000; Rothman 2002). This is an important difference, in that bias and confounding become major considerations in interpreting the results. In an RCT, randomization minimizes the chance that differences in outcomes between the experimental and control groups are the result of preexisting differences in the subjects in the two groups. In a cohort study, preexisting differences between exposed and nonexposed subjects may influence the subsequent risk of disease.

Confounding is especially problematic in studies of medication side effects, where the prescription of particular medications may be influenced by the preexisting medical conditions of patients, which is called *confounding by indication* (Rothman 2002). For example, because olanzapine causes more weight gain than risperidone (Lawrie and McIntosh 2002), risperidone might be prescribed more frequently to patients predisposed to obesity or diabetes, thus confounding any observed relationship between medication and risk of diabetes. Confounding can be accounted for in the statistical analysis of the data, but only to the extent that the confounding factors have been identified and measured. Statistical methods to control for confounding in cohort studies include stratified analysis and Poisson regression, details of which can be found in more advanced texts (Kelsey et al. 1996; Rothman and Greenland 1998).

There are also several potential biases in cohort studies that could be problematic. These include differential misclassification as to disease and exposure status, losses to follow-up, and nonparticipation. Good discussions of biases and confounding in cohort studies can be found in most standard epidemiology texts (Gordis 2000; Hennekens et al. 1987; Rothman 2002).

Measures of Risk

In a cohort study, incidence rates are calculated for exposed and nonexposed groups. Several measures of risk can then be calculated, including absolute risk, RR, odds ratio (OR), and risk difference (RD) (Table 11–3). RD in a cohort study is analogous to ARI in an RCT, and NNH may be calculated as 1/RD. RR and OR are measures of the strength of association between an exposure and outcome, whereas RD and NNH are better measures of the potential for prevention (Gordis 2000; Sackett et al. 2000).

TABLE 11–3. Measures of risk: cohort study

	Exposed group	Nonexposed group
Harm	A	B
No harm	C	D
Total	A+C	B+D

Incidence in exposed group=A/(A+C)
Incidence in nonexposed group=B/(B+D)
Relative risk=incidence in exposed group/incidence in nonexposed group
Risk difference=incidence in exposed group−incidence in nonexposed group
Number needed to harm=1/RD

Case-Control Studies

The other major study design is the case-control study. In a case-control study, patients with a particular disease (cases) are compared with patients without the disease in question (i.e., controls) with regard to exposures and other characteristics. Case-control studies are especially useful in studying rare diseases, because cohort studies involve studying very large cohorts for extended periods of time for meaningful numbers of cases of rare diseases to occur.

By their nature, however, case-control studies are more susceptible to bias than are cohort studies and therefore rank lower on the hierarchy of evidence (Table 4–1, Chapter 4). Two of the major sources of bias relate to the choice of control subjects (selection bias) and the fact that information about exposure is gathered after the onset of disease (recall or information bias).

Choice of Controls

In a case-control study, control subjects must be chosen from the same source population as the cases, because control subjects are used to estimate the prevalence of exposure to a risk factor in the population from which the cases are drawn (Lewis and Pelosi 1990; Rothman 2002). In other words, if cases are identified from a particular clinic or hospital, control subjects must be individuals who similarly have obtained treatment from that same clinic or hospital if they have the disease in question. In many cases, there are multiple medical providers serving a geographic area; therefore, hypothetical source population is not defined by specific geographic boundaries, but rather by referral or care-seeking patterns.

Several potential sources of control subjects can be used, and each has its own advantages and limitations (Hennekens et al. 1987; Rothman 2002). Sources include general population controls, including those identified by random-digit dialing; hospital or clinic controls; and friends, relatives, or neighbors of the cases.

Although it is often convenient to draw control subjects from the same hospital or clinic as the cases, because they clearly come from the same source population, there are also some important limitations related to exposure status. In particular, hospitalized control subjects differ from individuals without disease in the frequency of exposures that are associated with the control subjects' own diseases. For example, if an investigator is interested in conducting a case-control study to determine if alcohol use is a risk factor for deliberate self-harm, it might be tempting to identify both cases and controls from hospital emergency room patients. Because alcohol use is also a risk factor for accidental injuries, violence, and a number of other diseases, obtaining control subjects from an emergency room overestimates the use of alcohol in the population without disease and hence tends to obscure any association with self-harm.

The "nested" case-control study, which minimizes some of the issues in control selection, is a variant of the case-control study (Gordis 2000; Rothman 2002). In a nested case-control study, both cases and controls come from a previously assembled cohort; therefore, there is an enumerated source population from which to select the controls. An example of a nested case-control study is that of Koro et al. (2002), in which cases of patients with diabetes and control subjects were obtained from a cohort of over 19,000 patients with schizophrenia who were in a general practice research database. The authors then assessed prior use of antipsychotic medication by the patients with diabetes and the control subjects and were able to demonstrate that olanzapine use increased the risk of developing diabetes.

Recall Bias

The other serious concern in case-control studies is recall bias. Individuals with a disease often search their memories for possible causes and are therefore more likely to selectively recall exposures than are control subjects; this is sometimes called *rumination bias* or *effort after meaning* (Creed 1993; Gordis 2000). In psychiatric epidemiology, there is the added problem that the psychiatric illness may selectively influence memory (Lewis and Pelosi 1990). For example, patients with depression are more apt to remember negative life experiences than are nondepressed individuals (Creed 1993; Lloyd and Lishman 1976).

Reverse Causality

An additional problem with case-control studies is that of reverse causality (Creed 1993; Lewis and Pelosi 1990). In a case-control study, cases are asked about exposures prior to the onset of illness. For some psychiatric illnesses, it is difficult to determine the exact onset of illness. In addition, it may be difficult to determine retrospectively if an event, such as marital discord, truly preceded the onset of illness or if it was a consequence of the illness. This has been a

particular problem with life events and depression research (Cooper and Paykel 1993; Creed 1993).

Odds Ratio

The odds ratio is the measure of risk in a case-control study (Table 11–4). For rare diseases, the OR approximates RR from a cohort study. Measures of absolute risk cannot be directly estimated from a case-control study. It is possible, however, to estimate NNH from the OR, if the incidence of disease in the unexposed population is known (Bjerre and LeLorier 2000). This method is described in Appendix B.

Causality in Epidemiologic Studies

There are several explanations for an association between an exposure and a disease in an epidemiologic study (Table 11–5), including bias, chance, confounding, and causality (Fletcher et al. 1996; Hennekens et al. 1987).

Bias

Bias refers to a systematic error that results in an incorrect estimate of the risk of disease associated with an exposure. Bias may occur as a result of the process of selecting subjects (selection bias) or from gathering information on exposures (information or recall bias). As noted earlier in this chapter, case-control studies are more susceptible to selection and information biases than are prospective cohort studies.

Chance

Chance is always a possible explanation for an observed association between an exposure and a disease. The likelihood that chance alone is responsible is assessed through tests of statistical significance (*P* values) and confidence intervals (see Chapter 5).

Confounding

Confounding involves the possibility that differences in subjects (other than the exposure under investigation) are responsible for the observed association between the exposure and the disease. For example, if women in a retirement community are, on average, older than men, an observed increased risk of Alzheimer's disease in women could be the result of an association with age rather than with female sex. In other words, age is a confounding factor. If there are potential confounding factors that have been measured in a study, the statistical analysis of the data can take them into account, either through stratified analysis or regression techniques (i.e., Poisson regression for a cohort study or logistic regression for a case-control study) (Kelsey et al. 1996; Rothman and Greenland 1998).

TABLE 11–4. Measure of risk: case-control study

	Cases	Controls
Exposed	A	B
Nonexposed	C	D

Odds ratio = (A/C)/(B/D) = (AD)/(BC)

Causality

If bias, chance, and confounding are not the explanations for an association between exposure and disease, a causal relationship is likely. Bradford Hill (1965) described a variety of evidence that would support a causal relationship. Table 11–6 presents some of the criteria for causation suggested by Bradford Hill (1965) and others (Elwood 1998; Fletcher et al. 1996; Gordis 2000; Hennekens et al. 1987; Turcotte et al. 2001). Although Rothman (2002) has criticized such "checklists" as not reflecting more sophisticated notions of causality, many other epidemiologists continue to find them useful.

Critical Appraisal Guides for Studies of Etiology or Harm

Tables 11–7 through 11–9 are the critical appraisal guides for studies of etiology or harm. Separate guides are provided for RCT, cohort study, and case-control study, the three major study designs.

TABLE 11–5. Explanations for associations between exposure and disease in epidemiologic studies

Bias
Selection bias
Information bias
Chance
Confounding
Causality

TABLE 11–6. Suggested criteria for causality in epidemiologic studies

Criterion	Description
Temporal relationship	Exposure precedes disease
Strength of association	Large relative risk or odds ratio
Dose–response relationship	Increasing exposure increases risk
Reversibility	Reducing exposure decreases risk
Consistency	Similar findings from other studies in different populations
Biological plausibility	Consistent with pharmacological or toxicological data
Analogy	Relationship established for similar cause and disease
Elimination of other explanations	Not the result of confounding or bias
Specificity	One cause, one effect

Source. Adapted from Fletcher et al. 1996; Gordis 2000; Hennekens et al. 1987; Sackett et al. 2000.

Randomized Controlled Trial Reporting Harm

Is the Study Valid?

Before reviewing the results of the RCT, you should first ensure that the study is valid. The criteria for doing this are similar to those for appraising studies designed to evaluate the efficacy of a therapy (Table 5–6, Chapter 5). However, there are some additional considerations for a study of adverse effects.

First, were the adverse effects assessed in a systematic way (as opposed to self-reporting)? Some side effects (e.g., sexual side effects) may not be spontaneously reported because of embarrassment. Not asking about a specific side effect may lead to misleadingly low rates.

Second, was the size of the study sufficient to detect meaningful differences in rates of the side effects of interest? RCTs lack sufficient power to detect rare side effects; however, if a particular side effect (e.g., weight gain) is of interest, the study should have sufficient power to detect meaningful differences.

Are the Results Important?

In an RCT that reports adverse effects, NNH is the measure of interest. In addition, the severity of side effects, not only the frequency, is important.

Can I Apply the Results to My Patient?

As noted in Chapter 5, this question can be reframed as: "Is the pathobiology of my patient so different from that of the study patients that the results cannot apply?" Here the considerations regarding side effects become more complicated than with the assessment of beneficial effects. There are often significant differences in the risk of adverse effects, depending on patient characteristics. To take an obvious example, men and nonpregnant women are not at risk for the teratogenic side effects of a medication. Beyond this, however, there are a variety of physical illnesses that could be exacerbated by a medication or that could predispose a patient to a particular side effect. In addition, there are age-related changes in drug metabolism, and members of particular ethnic groups may be at increased risk of specific medication side effects, partly because of allelic variation in genes coding for drug-metabolizing enzymes (Cookson et al. 2002; Ruiz 2000). Thus, some judgment is required in deciding whether your patient is enough like those in the study that the results apply and whether your patient is at increased or decreased risk. In general, your patient will probably be similar enough to those studied that the results will apply, although the magnitude of the risk may differ.

In assessing the relative risks and benefits of treatment, the issue is one of comparing NNT as a measure of treatment effectiveness with NNH as a measure of treatment side effects. This can certainly be done informally by simply comparing NNT and NNH, together with a subjective assessment of the relative value the patient places on the benefit versus the risk. A more formal quantitative method of doing this has been described that involves calculating a statistic, the likelihood of being helped or harmed

TABLE 11–7. Critical appraisal guide for randomized controlled trials reporting harm

Is the study valid?
Is it a randomized controlled trial?
Was the randomization list concealed?
Were subjects and clinicians blinded to treatment being administered?
Were side effects assessed appropriately?
Was the trial of sufficient duration to detect the side effects of interest?
Were all subjects enrolled in the trial accounted for?
Were subjects analyzed in the groups to which they were assigned?
Despite randomization, were there clinically important differences between groups at the start of the trial?
Aside from the experimental treatment, were the groups treated equally?
Are the results important?
How large was the treatment effect (number needed to harm)?
How precise are the results (width of confidence intervals)?
Can I apply the results to my patient?
Is my patient too different from patients in the study?
How do the benefits and risks of treatment compare for my patient?
How does my patient value these benefits and risks?
Do the benefits outweigh the harms (likelihood of being helped or harmed)?

Source. Based in part on the criteria of Levine et al. 2002; Sackett et al. 2000.

(LHH) (Guyatt et al. 2002; Sackett et al. 2000). As a first approximation:

$$\text{LHH} = \frac{1/\text{NNT}}{1/\text{NNH}} = \frac{\text{NNH}}{\text{NNT}}$$

with values of LHH>1 indicating that the benefit outweighs the harm. This crude calculation weights benefits and side effects equally. A more sophisticated calculation, taking into account patients' relative preferences for side effects versus treatment effects, (as well as patient-specific NNTs and NNHs) is given in Appendix B.

Cohort Studies of Etiology Or Harm

A critical appraisal guide for a cohort study of etiology or harm is given in Table 11–8.

Is the Study Valid?

Several questions should be asked to determine the validity of the study. First, how were the subjects selected? The cohort should consist of individuals who are initially free of the disease under investigation, but who are potentially at risk for developing the disease. In addition, the exposed and nonexposed groups should be selected in such a way as to avoid major differences other than exposure status.

Was the cohort large enough and was the follow-up period long enough? For rare outcomes, large numbers of individuals must often be studied for prolonged periods of time to detect statistically significant differences in risk.

How was exposure measured? If significant errors are made in classifying individuals as either "exposed" or not, the results will tend to be biased. Random misclassification will tend to bias the observed RR toward 1.0, whereas differential misclassification will result in either an overestimate or underestimate of the actual RR (Hennekens et al. 1987).

Were exposed and unexposed individuals assessed for outcomes with the same intensity and were outcomes assessed blindly (i.e., without knowledge of exposure status)? If not, any differences could be the result of measurement bias, not true differences (Fletcher et al. 1996).

Similarly, were there differences in dropout rates between the two groups? Because the outcomes of patients who drop out may differ from the outcomes of patients completing the study, differences in dropout rates between the two groups may lead to biased estimates of risk (Hennekens et al. 1987).

Were the exposed and nonexposed groups in fact similar, except for exposure? If not, then confounding may be responsible for any difference in risk that is observed. If there were differences in potential confounding factors, appropriate statistical techniques (stratification or multivariate techniques) should have been used in the data analysis.

TABLE 11–8. Critical appraisal guide for cohort studies

Is the study valid?
Were the exposed and nonexposed groups similar (other than exposure) to the risk factor at the onset of the study?
If there were differences between groups at the start of the trial, did the statistical analysis take the differences into account?
Was the follow-up of the cohort sufficiently long for outcomes to develop?
Were outcomes measured in the same way in both groups?
Were there significant differences in losses to follow-up in the two groups?
Were any of the criteria for causality met (Table 11–6)?
Are the results important?
How strong is the association (relative risk)?
How large is the absolute increase in risk (risk difference)?
Can I apply the results to my patient?
Is my patient too different from patients in the study?
How do the benefits and risks of treatment compare for my patient?
How does my patient value these benefits and risks?
Do the benefits outweigh the harms (likelihood of being helped or harmed)?

Source. Based in part on the criteria of Elwood 1998; Levine et al. 2002; Sackett et al. 2000.

Were any of the criteria for causality (Table 11–6) met? If so, the likelihood that the results were not the result of chance, bias, or confounding is increased.

Are the Results Important?

In a cohort study, RR measures the strength of association between the exposure and outcome, whereas NNH is a better measure of the potential harm to an individual patient (Sackett et al. 2000).

Can I Apply the Results to My Patient?

As with RCTs, this question can be reframed in the following way: "Is the biology of my patient so different from that of the study patients that the results cannot apply?"

If the clinical question is one that concerns the etiology of a disorder, the results will generally be applicable, assuming that your patient could be at risk for the outcome and bearing in mind that some outcomes are limited to patients who are of a particular age, sex, or childbearing status. With regard to etiology, the assumptions are that RR will apply to all individuals and that risks are generally multiplicative. In certain circumstances, however, interactions between risk factors that could either increase or decrease a particular patient's risk beyond what is expected by the RR associated with the individual risk factors are possible (Gordis 2000; Rothman 2002).

In assessing the absolute risk to an individual patient, however, NNH is used as the measure of risk. Unlike RR, this measure is sensitive to the individual's baseline risk and may need to be individualized.

As with an RCT, in assessing the relative risks and benefits of treatment, the issue is one of comparing NNT (usually derived from an RCT) with NNH from the cohort study, and the same considerations apply as discussed for LHH above.

Case-Control Study

A critical appraisal guide for a case-control study is given in Table 11–9.

Is the Study Valid?

The first question is how the cases were chosen. The cases should be representative of patients with the disease. In addition, they should be incident (newly diagnosed) cases, because a study of prevalent cases may reveal more about risk factors for chronicity than about the etiology of the disorder.

The next question is how the control subjects were chosen. In a case-control study, control subjects must be chosen from the same source population as the cases, because control subjects are used to estimate the prevalence of exposure to a risk factor in the population from which the cases are drawn (Lewis and Pelosi 1990; Rothman 2002). In other words, if the cases in a study were identified from a particular clinic or hospital, would the control subjects have obtained treatment from that same clinic or hospital if they had the disease in question? Control subjects can include general population controls; hospital or clinic controls; and friends, relatives, or neighbors of the cases. If more than one

TABLE 11–9. Critical appraisal guide for case-control studies

Is the study valid?
Did the method of selection of cases and control subjects introduce bias?
Were outcomes measured in the same way in both groups, independent of disease status?
Were confounding factors identified and dealt with in the analysis?
Were any of the criteria for causality met (Table 11–6)?
Are the results important?
How strong is the association (odds ratio)?
Can I apply the results to my patient?
Is my patient too different from patients in the study?
What is the number needed to harm for my patient?
How do the benefits and risks of treatment compare for my patient?
How does my patient value these benefits and risks?
Do the benefits outweigh the harms (likelihood of being helped or harmed)?

Source. Based in part on the criteria of Fletcher et al. 1996; Levine et al. 2002; Mant 1999; Sackett et al. 2000.

control group was used in a particular study, were the results similar? If so, it is less likely that the results were the result of selection bias.

The other serious concern in case-control studies is information or recall bias. Exposure history should be gathered without knowledge of whether the individual is a case or a control subject. In addition, using more than one information source may increase the validity of the information being collected, because individuals with a disease often search their memories for possible causes, and psychiatric illnesses may selectively influence memory.

Were the cases and the control subjects generally similar except for exposure? If not, confounding may be responsible for any difference in the amount of risk that is observed. If there were differences in potential confounding factors, appropriate statistical techniques (stratification or multivariate techniques) should have been used in the data analysis.

Were any of the criteria for causality (Table 11–6) met? If so, the likelihood that the results were not the result of chance, bias, or confounding is increased.

Are the Results Important?

In a case-control study, the OR measures the strength of association between the exposure and the outcome. Measures of absolute risk cannot be determined from a case-control study alone.

Can I Apply the Results to My Patient?

As in a cohort study, if the clinical question concerns the etiology of a disorder, the results will generally be applicable, assuming your patient could be at risk for the outcome. The assumptions are that the OR will apply to all individuals and that risks are multiplicative. There may be interactions between risk factors that may either increase or decrease a particular patient's risk beyond what would be expected from the OR associated with the individual risk factors (Gordis 2000; Rothman 2002).

Because measures of absolute risk cannot be directly estimated from a case-control study, the study itself will not yield NNH. It is possible, however, to estimate NNH from the OR if the probability of disease in the patient (usually estimated from the incidence of disease in the population) is known (Bjerre and LeLorier 2000). A method for calculating this estimation is given in Appendix B. As in an RCT or cohort study, the issue in assessing the relative risks and benefits of treatment becomes one of comparing NNT (usually derived from an RCT) with NNH (estimated from the OR and population incidence data).

References

Bjerre LM, LeLorier J: Expressing the magnitude of adverse effects in case-control studies: "the number of patients needed to be treated for one additional patient to be harmed." BMJ 320:503–506, 2000

Bradford Hill AB: The environment and disease: association and causation. Proc R Soc Med 58:295–300, 1965

Brewer T, Colditz GA: Postmarketing surveillance and adverse drug reactions: current perspectives and future needs. JAMA 281:824–829, 1999

Cookson J, Taylor D, Katona C: The Use of Drugs in Psychiatry: The Evidence from Psychopharmacology, 5th Edition. London, Gaskell, 2002

Cooper Z, Paykel ES: Social factors in the onset and maintenance of depression, in Principles of Social Psychiatry. Edited by Bhugra D, Leff J. London, Blackwell Scientific, 1993, pp 99–121

Creed F: Life events, in Principles of Social Psychiatry. Edited by Bhugra D, Leff J. London, Blackwell Scientific, 1993, pp 144–161

Elwood M: Critical Appraisal of Epidemiological Studies and Clinical Trials, 2nd Edition. Oxford, UK, Oxford University Press, 1998

Fletcher RH, Fletcher SW, Wagner EH: Clinical Epidemiology: The Essentials, 3rd Edition. Baltimore, MD, Williams & Wilkins, 1996

Gordis L: Epidemiology, 2nd Edition. Philadelphia, PA, WB Saunders, 2000

Gunnell D, Harrison G, Rasmussen F, et al: Associations between premorbid intellectual performance, early life exposures and early onset schizophrenia: cohort study. Br J Psychiatry 181:298–305, 2002

Guyatt G, Straus S, McAlister F, et al: Moving from evidence to action: incorporating patient values, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 567–582

Hennekens CH, Buring JE, Mayrent SL: Epidemiology in Medicine. Boston, MA, Little, Brown, 1987

Ioannidis JPA, Lau J: Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. JAMA 285:437–443, 2001

Jones TC: Call for a new approach to the process of clinical trials and drug registration. BMJ 322:920–923, 2001

Kelsey JL, Whittemore AS, Evans AS, et al (eds): Methods in Observational Epidemiology, 2nd Edition. New York, Oxford University Press, 1996

Ko DT, Hebert PR, Coffey CS, et al: Beta-blocker therapy and symptoms of depression, fatigue, and sexual dysfunction. JAMA 288:351–357, 2002

Koro CE, Fedder DO, L'Italien GJ, et al: Assessment of independent effect of olanzapine and risperidone on risk of diabetes among patients with schizophrenia: population based nested case-control study. BMJ 325:243–245, 2002

Lawrie S, McIntosh A: Schizophrenia. Clin Evid 7:920–944, 2002

Levine M, Haslam D, Walter S, et al: Harm, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 81–100

Lewis G, Pelosi AJ: The case-control study in psychiatry. Br J Psychiatry 157:197–207, 1990

Lloyd G, Lishman A: The effect of depression on speed of recall of pleasant and unpleasant experiences. Psychol Med 5:173–190, 1976

Mant J: Case-control studies, in Evidence-Based Practice: A Primer for Health Care Professionals. Edited by Dawes M, Davies P, Gray A, et al. Edinburgh, Churchill Livingstone, 1999, pp 73–84

Pirmohamed M, Breckenridge AM, Kitteringham NR, et al: Adverse drug reactions. BMJ 316:1295–1298, 1998

Rothman KJ: Epidemiology: An Introduction. New York, Oxford University Press, 2002

Rothman KJ, Greenland S (eds): Modern Epidemiology, 2nd Edition. Philadelphia, PA, Lippincott-Raven, 1998

Ruiz P (ed): Ethnicity and Psychopharmacology (Review of Psychiatry Series, Vol 19, No 4). Washington, DC, American Psychiatric Press, 2000

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Takeshita J, Masaki K, Ahmed I, et al: Are depressive symptoms a risk factor for mortality in elderly Japanese American men? The Honolulu-Asia Aging Study. Am J Psychiatry 159:1127–1132, 2002

Turcotte K, Raina P, Moyer VA: Above all, do no harm. West J Med 174: 325–329, 2001

12

Studies of Prognosis

Gregory E. Gray, M.D., Ph.D.

Questions about prognosis are frequently raised by patients and by their families. Such questions take many forms. When will I get better? Will I be completely well? What are the chances of the disease recurring? Such questions are studied by following groups of patients over time (i.e., in a cohort study). Some such studies are purely descriptive, whereas other studies attempt to find prognostic factors that predict a good or bad outcome. The latter types are similar to the cohort studies of risk described in Chapter 11 and are subject to the same design and analysis issues. This chapter focuses on descriptive studies of prognosis and briefly reviews some of the issues in their design and analysis before discussing their critical appraisal.

Selecting Patients

The selection of patients in a cohort study of prognosis is of considerable importance. The two key issues are the populations from which the subjects are obtained and whether they are newly diagnosed cases or patients currently in treatment.

Source of Patients

One can think of several populations of patients with the same illness: patients in the general population who are in treatment with non–mental health practitioners, those in outpatient treatment with mental health professionals, those in specialized mental health clinics (e.g., an academic mood disorders clinic), and those who are hospitalized for their conditions. In general, patients seen at academic medical centers have more severe disease, more comorbid conditions, and are more likely to be treatment resistant or chronically ill than are those in community settings (Cohen and Cohen 1984; Hulley et al. 2001; Randolph et al. 2002). We know far less about the prognosis of individuals in the community who do not seek treatment (Regier et al. 1998), but it is believed that they have a better prognosis and more social supports (Cohen and Cohen 1984).

An example of such a difference in prognosis comes from a study of the duration of major depressive episodes in the general population conducted by Spijker et al. (2002). In this study, cases were identified as part of a prospective epidemiological study in a community, so the authors were able to identify cases that did not seek treatment, as well as those that did. Median durations of the depressive episode were 3.0 months for patients who had no professional care, 4.5 months for patients treated in primary care settings, and 6.0 months for patients who entered the mental health system.

There is no "right" population to study, but the population studied affects the generalizability of the results (Gordis 2000; Randolph et al. 2002). Results obtained from a study of patients in an academic medical center may not apply to individuals in the community.

Incident Versus Prevalent Cases

The other major issue concerns whether to select new (incident) cases or existing (prevalent) cases. There is a temptation to begin with a cohort of patients already in treatment (a survival cohort), but

such a sample would contain an overrepresentation of chronic patients and hence lead to a biased estimate of prognosis (Cohen and Cohen 1984; Fletcher et al. 1996). Instead, an inception cohort (a group of people assembled near the onset of disease) should be studied (Fletcher et al. 1996).

A related issue is that of the starting point or *zero time* (Fletcher et al. 1996; Gordis 2000). This could either be at the onset of symptoms, when treatment was first begun, or when a diagnosis was made. Regardless of what is chosen as the zero time, it should be used consistently as the starting point (Fletcher et al. 1996).

Follow-up and Attrition

One of the major sources of bias in cohort studies of prognosis concerns patients lost to follow-up. If patients lost to follow-up differ in outcomes from patients who remain in the study, estimates of prognosis will be biased. This is called *migration bias* (Fletcher et al. 1996).

One way of dealing with losses to follow-up in the analysis is to perform a *best case/worst case* analysis, a form of sensitivity analysis in which outcome statistics are first calculated assuming all of those lost to follow-up did well, then recalculated assuming all of those lost to follow-up had a bad outcome (Fletcher et al. 1996; Sackett et al. 2000). Ideally, researchers should attempt to minimize losses to follow-up through periodic contact and other means (Cummings et al. 2001).

Outcomes and Data Analysis

There are two commonly used approaches for describing the prognosis of a cohort: reporting on outcomes at specified follow-up times and measuring time to an event.

Outcomes at Specific Follow-Up Times

Investigators will sometimes recontact a cohort at specified times and collect data for a variety of outcome measures. Such an approach was taken by Wiersma et al. (2000) when they used several measures of social disability to assess a cohort of patients with schizophrenia at 1, 2, and 15 years after diagnosis. The advantage of this approach is that considerable information can be collected at each follow-up interview. The disadvantage is that no information is gathered between the set follow-up intervals, and the losses that follow lead to data that are based on smaller and smaller numbers of subjects at each subsequent interview.

Time-to-Event Outcomes: Survival Analysis

The use of time-to-event as an outcome measure in studies of prognosis is more common than the use of specified follow-up times. The event may be a negative one (e.g., death, relapse, rehospitalization, or dropping out of treatment) or a positive one (e.g., recovery). Such data are analyzed using statistical techniques called *survival analysis* or *failure time analysis*.

Survival analysis acknowledges that patients may be lost to follow-up; however, the analysis does include these patients until such time as they are lost. A basic statistical assumption, however, is that the prognosis of patients lost to follow-up (i.e., censored) will be the same as for patients who continue in the study (Bland and Altman 1998; Gordis 2000). Survival curves, such as the one in Figure 12–1, are calculated using the Kaplan-Meier method, the details of which are given in standard biostatistics texts (e.g., Altman 1991; Pagano and Gauvreau 2000). The log-rank test can be used to test for differences in survival times between subgroups (Altman 1991; Peto et al. 1977).

Critical Appraisal Guide for Studies of Prognosis

Guidelines for appraising cohort studies of prognosis are given in Table 12–1. As with other types of studies, validity should be assessed before considering the results.

Is the Study Valid?

The study should be based on an inception (incidence) cohort. If it is based on a survival cohort of patients at various stages of illness, it is simply not a valid study of prognosis.

The next question concerns follow-up losses. Sackett et al. (2000) have suggested using the "5 and 20" guideline, where <5% of patients lost to follow-up is probably not significant, whereas >20% lost to follow-up seriously threatens the validity of the study.

Finally, it is important that study outcomes be assessed in a uniform way.

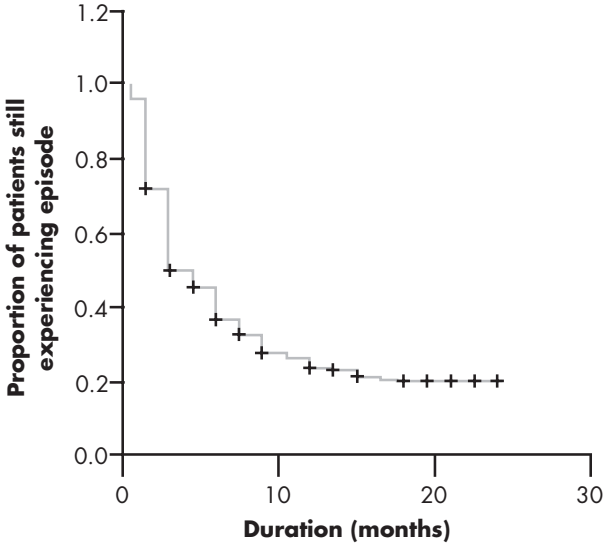


FIGURE 12–1. Survival curve of a cohort with newly originated major depression in the community.

Source. Reprinted from Spijker J, DeGraaf R, Bijl RV, et al: "Duration of Major Depressive Episodes in the General Population: Results From The Netherlands Mental Health Survey and Incidence Study (NEMESIS)." *British Journal of Psychiatry* 181:208–213, 2002. Copyright 2002 Royal College of Psychiatrists. Used with permission.

What Are the Results?

If measures other than time-to-event data are used, results should be presented with confidence intervals. Time-to-event data are generally presented in the form of a survival curve, often omitting confidence intervals for the curve itself but providing some information in the text on precision of the estimates.

Can I Apply the Results to My Patient?

As noted above, the prognosis of a community sample may differ significantly from that obtained from various treatment settings. Whenever possible, an attempt should be made to find a study reporting on a population from a setting similar to that of your patient. If this is not possible, the results may still be usable, keeping in mind the general rule that cohorts assembled from an academic medical center or specialty mental health program will often contain a greater number of poor-prognosis patients than will cohorts assembled from primary care or community settings.

TABLE 12–1. Critical appraisal guide for studies of prognosis

Is the study valid?
Is it a cohort study based on an inception cohort?
Was the follow-up relatively complete (>80% and preferably >95%)?
Were outcomes assessed in a uniform, unbiased manner?
What are the results?
What are the outcome data at various points in time?
Is there a survival curve?
How precise are the estimates?
Can I apply the results to my patient?
Were the patients similar to my patients in diagnosis and comorbidity?
Were the patients derived from a similar treatment setting?
Was the management similar to that in my practice?

Source. Based in part on data from Randolph et al. 2002; Sackett et al. 2000.

A secondary consideration has to do with comorbidity and other patient attributes that may affect prognosis. Again, there are no hard and fast rules, but patients with comorbid conditions frequently have a worse prognosis than do patients without comorbid conditions. Once again, although this does not invalidate a study, some mental adjustment of the results must be made in applying them to your patient.

References

Altman DG: Practical Statistics for Medical Research. Boca Raton, FL, Chapman & Hall, 1991

Bland JM, Altman DG: Survival probabilities (the Kaplan-Meier method). BMJ 317:1572, 1998

Cohen P, Cohen J: The clinician's illusion. Arch Gen Psychiatry 41:1178–1182, 1984

Cummings SR, Newman TB, Hulley SB: Designing an observational study: cohort studies, in Designing Clinical Research: An Epidemiologic Approach, 2nd Edition. Edited by Hulley SB, Cummings SR. Philadelphia, PA, Lippincott Williams & Wilkins, 2001, pp 95–105

Fletcher RH, Fletcher SW, Wagner EH: Clinical Epidemiology: The Essentials, 3rd Edition. Baltimore, MD, Williams & Wilkins, 1996

Gordis L: Epidemiology, 2nd Edition. Philadelphia, PA, WB Saunders, 2000

Hulley SB, Newman TB, Cummings SR: Choosing the study subjects: specification, sampling, and recruitment, in Designing Clinical Research: An Epidemiologic Approach, 2nd Edition. Edited by Hulley SB, Cummings SR. Philadelphia, PA, Lippincott Williams & Wilkins, 2001, pp 25–35

Pagano M, Gauvreau K: Principles of Biostatistics, 2nd Edition. Pacific Grove, CA, Duxbury, 2000

Peto R, Pike MC, Armitage P, et al: Design and analysis of randomized clinical trials requiring prolonged observation of each patient, II: analysis and examples. Br J Cancer 35:1–39, 1977

Randolph A, Bucher H, Richardson WS, et al: Prognosis, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt G, Rennie D. Chicago, IL, AMA Press, 2002, pp 141–154

Regier DA, Kaelber CT, Rae DS, et al: Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. Arch Gen Psychiatry 55:109–115, 1998

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Spijker J, DeGraaf R, Bijl RV, et al: Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry 181:208–213, 2002

Wiersma D, Wanderling J, Dragomirecka E, et al: Social disability in schizophrenia: its development and prediction over 15 years in incidence cohorts in six European cities. Psychol Med 30:1155–1167, 2000

13

Evaluating Your Performance of Evidence-Based Medicine

Gregory E. Gray, M.D., Ph.D.

Psychiatry residents are expected to acquire positive attitudes toward lifelong learning, and more specifically, according to the Accreditation Council for Graduate Medical Education (ACGME): "The physician shall recognize limitations in his or her own knowledge base and clinical skills, and understand and address the need for lifelong learning." In this chapter, the focus is on evaluating evidence-based medicine (EBM) skills. In the next chapter, we focus on the broader issues of evaluating one's practice.

In the 5-step EBM model (see Gray, Chapter 2 in this volume), the final step is assessing the outcome. In ordinary clinical practice, we assess whether a particular treatment worked or whether a diagnostic test provided helpful information. In the EBM model, assessment has the added component of evaluating the clinician's performance in the EBM process.

The model for such an evaluation, as suggested by Sackett et al. (2000) and Wolf (2000), focuses on skills in performing each of the five steps in the EBM process (Table 13–1). A more in-depth discussion of the evaluation of clinical skills in general—and of EBM skills in particular—can be found in the works of Sackett et al. (1991, 2000).

Formulating Clinical Questions

The first necessary skill is formulating the 4-part clinical questions (see Gray, Chapter 3 in this volume). Such a skill is important because it leads to a more efficient search strategy (Cabell et al. 2001), yet it is still one with which physicians may have difficulty (Ely et al. 2002).

The first question to ask is whether the clinician understands the concept of the 4-part PICO clinical question (i.e., patient/problem, intervention, comparison, and outcome, as described in Chapter 3) and whether he or she was able to formulate such questions before beginning the most recent search for information. The next question to ask is whether the clinician routinely formulates such questions in his or her clinical practice.

Clinical questions arise every day in clinical practice, but we often do not have the ability to answer them on the spot (Del Mar and Glasziou 2001). To incorporate EBM into daily practice requires that we have a way of keeping track of the most important questions so that they can be answered at a later time.

Searching for Answers

The next step in the process is searching for an answer to the clinical question. In evaluating a particular search, barriers to finding the information should be identified. Common barriers include lack of time and information resources, as well as the search strategy itself (Cabell et al. 2001; Craig et al. 2001; Ely et al. 2002; McColl et al. 1998). Furthermore, many practitioners either lack awareness of the Cochrane Database of Systematic Reviews and other resources or do not use them, even if available (Kerse et al. 2001; McColl et al. 1998).

TABLE 13–1. Self-evaluation of evidence-based medicine skills

Formulating clinical questions
Do I understand what a 4-part PICO clinical question is?
Was I able to formulate a 4-part PICO question this time?
Do I routinely keep track of clinical questions that arise in my practice and attempt to find answers to them?
Searching for answers
Do I know the common databases and how to search them?
Did this particular search go well or could I have been more efficient?
Do I routinely attempt to answer clinical questions through searches or do I still rely primarily on textbooks?
Appraising the evidence
Do I understand how to critically appraise research articles and other evidence?
Was I able to apply the critical appraisal guide in this case, including considering my patient's individual risks, needs, and preferences?
Are my critical appraisal skills improving over time?
Applying the evidence
Do I incorporate the evidence that I have appraised (and have found to be valid) into my clinical practice?
What proportion of my clinical decisions is based on current best evidence?
Evaluating the results
Do I routinely evaluate my evidence-based medicine skills?
Are there particular aspects of the evidence-based medicine process that I need to review?

Note. PICO=Patient/Problem, Intervention, Comparison, and Outcome.

In evaluating one's own performance, the first question to ask is whether you have an understanding of the basic search process (as outlined by Gray and Taylor in Chapter 4 in this volume), including the roles of the various databases that are available. The second question relates to the ability to perform the particular search and whether another search strategy would have proved more efficient. The final question to ask is whether you are routinely searching for answers in appropriate databases in your clinical practice rather than relying on textbooks.

Appraising the Evidence

Critical appraisal skills are the most commonly taught aspect of EBM (Green 1999), yet clinicians often remain unfamiliar with many of the basic concepts (LeClair et al. 1999; McColl et al. 1998; Young et al. 2002). Such skills, however, are associated with the ability and willingness to apply the results of systematic reviews in clinical practice (Doust and Silagy 2000).

In evaluating yourself, the first question to ask is whether you have an understanding of the methods of critical appraisal, including the use of critical appraisal worksheets. (To review this information, see Chapters 5–11.) The next question to ask is whether you had any difficulties in appraising the evidence from this particular search. This also involves considering whether the evidence was applicable to your patient, being able to individualize the results for your patient, and assessing whether it was consistent with patient needs and preferences. Finally, you should ask whether your appraisal skills are improving over time as a result of practice.

Applying the Evidence

The fourth step involves applying the evidence to a particular patient. This is when clinicians often falter (Wolf 2000). The most important question to ask here is whether one's practice is becoming more evidence based. This can be done informally; for example, as the clinician sees patients, he or she should ask whether the treatment provided, diagnostic test ordered, and so forth were supported by good evidence. This may encourage clinicians to conduct searches for some common treatments that are being used, and the results may surprise them. It can also be done in a more formal way, as part of a quality improvement project (Baker and Grol 2001).

Evaluating the Results

Finally, clinicians should reflect on their own performance of EBM. In Chapter 14, Taylor addresses the issue of evaluating one's practice in more detail,

and in Chapter 15, Gray and Taylor provide resources to be used in teaching and learning EBM. Remember that skill in EBM is like any other skill: it gets easier with practice.

Learning Evidence-Based Medicine

Clinicians who need more training in EBM have several options to increase their knowledge and skills.

Books and Journals

This book covers the basics of EBM (and evidence-based psychiatric practice) as applied to psychiatric practice. Other excellent EBM texts, each with a slightly different orientation, provide additional information about various aspects of the EBM process. These include books by Dawes et al. (1999), Greenhalgh (2001), Guyatt and Rennie (2002a, 2002b), and Sackett et al. (2000). The texts by Dawes et al. (1999), Greenhalgh (2001), and Guyatt and Rennie (2002a) are similar in scope to the current volume, whereas the texts of Guyatt and Rennie (2002b), Sackett et al. (2000), and Straus et al. (2005) provide a more in-depth coverage of topics. Although the texts are more oriented toward internists or family practitioners, they can profitably be read by anyone who wants to apply EBM to psychiatric practice.

In addition to books specifically about EBM, books about the basic sciences behind EBM (i.e., clinical epidemiology and biostatistics) may be helpful. The best starting point for learning more about clinical epidemiology is the brief clinically oriented text by Fletcher et al. (1996); Gordis (2000) has written a useful text that is slightly more detailed. Among the many biostatistics texts, the monograph by Daly and Bourke (2000) stands out for its clarity and clinical relevance. Several journals also regularly include useful overviews of EBM concepts, including the *ACP Journal Club*, *Evidence-Based Mental Health*, and the general medical journals *BMJ* and *JAMA*.

Internet Resources

In addition to print resources, several online resources may be useful in learning more about various aspects of EBM. Some general resources are listed in Table 15–2 in Chapter 15 of this volume. A search of the Internet identifies numerous Internet-based courses as well.

Courses

Most medical libraries regularly offer courses about MEDLINE and other databases. In addition, medical librarians can be quite helpful in teaching you how to improve your search techniques.

Also, several regularly offered courses on EBM use active small-group teaching strategies. Many of these can be found by searching the Web sites listed in Table 4–2 in Chapter 4 of this volume. The American Psychiatric Association has also offered 1-day workshops at its annual meetings.

References

Baker R, Grol R: How to assess the effectiveness of applying the evidence, in Evidence-Based Practice in Primary Care, 2nd Edition. Edited by Silagy C, Haines A. London, BMJ Books, 2001, pp 83–97

Cabell CH, Schardt C, Sanders L, et al: Resident utilization of information technology: a randomized trial of clinical question formation. J Gen Intern Med 16:838–844, 2001

Craig JC, Irwig LM, Stockler MR: Evidence-based medicine: useful tools for decision making. Med J Aust 174:248–253, 2001

Daly L, Bourke GF: Interpretation and Uses of Medical Statistics, 5th Edition. Hoboken, NJ, Wiley-Blackwell, 2000

Dawes M, Davies P, Gray A, et al (eds): Evidence-Based Practice: A Primer for Health Care Professionals. New York, Churchill Livingstone, 1999

Del Mar CB, Glasziou PP: Ways of using evidence-based medicine in general practice. Med J Aust 174:347–350, 2001

Doust JA, Silagy CA: Applying the results of a systematic review in general practice. Med J Aust 172:153–156, 2000

Ely JW, Osheroff JA, Ebell MH, et al: Obstacles to answering doctors' questions about patient care with evidence: qualitative study. BMJ 324:710–713, 2002

Fletcher RH, Fletcher SW, Wagner EH: Clinical Epidemiology: The Essentials, 3rd Edition. Baltimore, MD, Williams & Wilkins, 1996

Gordis L: Epidemiology, 2nd Edition. Philadelphia, PA, WB Saunders, 2000

Green ML: Graduate medical education training in clinical epidemiology, critical appraisal, and evidence-based medicine: a critical review of the curricula. Acad Med 74:686–694, 1999

Greenhalgh T: How to Read a Paper: The Basics of Evidence Based Medicine, 2nd Edition. London, BMJ Books, 2001

Guyatt G, Rennie D (eds): Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice. Chicago, IL, AMA Press, 2002a

Guyatt G, Rennie D (eds): Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Chicago, IL, AMA Press, 2002b

Kerse N, Arroll B, Lloyd T, et al: Evidence databases, the Internet, and general practitioners: the New Zealand story. N Z Med J 114:89–91, 2001

LeClair BM, Wagner PJ, Miller MD: A tool to evaluate self-efficacy in evidence-based medicine. Acad Med 74:597, 1999

McColl A, Smith H, White P, et al: General practitioners' perceptions of the route to evidence based medicine: a questionnaire survey. BMJ 316:361–365, 1998

Sackett DL, Haynes RB, Guyatt GH, et al: Clinical Epidemiology: A Basic Science for Clinical Medicine, 2nd Edition. Boston, MA, Little, Brown, 1991

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Straus SE, Richardson WS, Glasziou P, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 3rd Edition. Edinburgh, Churchill Livingstone, 2005

Wolf FM: Lessons to be learned from evidence-based medicine: practice and promise of evidence-based medicine and evidence-based education. Med Teach 22:251–259, 2000

Young JM, Glasziou P, Ward JE: General practitioners' self ratings of skills in evidence based medicine: validation study. BMJ 324:950–951, 2002

14

Evidence-Based Psychiatric Practice

C. Barr Taylor, M.D.

The goals of evidence-based psychiatric practice are to improve patient care and to help ensure that patients receive the most effective treatments available. In a broader sense, the goal for the clinician also should be to frequently assess and improve his or her clinical skills and acumen. The Accreditation Council for Graduate Medical Education now requires residents to learn how to use practice guidelines as part of the systems-based competency (Andrews and Burruss 2004). In recent decades, the knowledge base in all aspects of psychiatry has increased dramatically, with new findings of relevance to patient care reported frequently in the scientific literature. Modern clinicians need to enhance their psychotherapy skills and techniques and to be conversant with new developments in medicine, psychotherapy, and psychopharmacology that may affect their patients' care. Keeping up to date on this knowledge requires the ability both to find information relevant to a particular patient and to monitor information that may be relevant to one's overall practice. In this chapter, I consider how clinicians can focus on the quality of care of their practice.

Evaluating one's practice is more complicated than finding an answer to a particular therapeutic question. In this case, the clinician would want to know not only about the choice of therapies but also how well the quality of his or her work measures up to standards in the field or literature and the clinician's own personal standards. Learning how to evaluate and improve their practice as needed is one of the core competencies psychiatry residents are now expected to acquire in the course of training. Evaluating one's practice is also a core feature of lifelong learning. In this chapter, I discuss several approaches to doing so.

Evaluating One's Practice

Sir William Osler, one of the founders of modern medicine, encouraged clinicians to classify cases as "clear cases, doubtful cases and mistakes." He argued that "It is only by getting your cases grouped in this way that you can make any real progress in your post-collegiate education; only in this way you gain wisdom with experience" (Cushing 1940, p. 328). He also said that "It is a common error to think the more a physician sees the greater his experience and more he knows," the implication being that experience is not sufficient for lifelong learning. Yet from an evidence-based standpoint, what constitutes a "clear case, doubtful case, and mistake"? A clinician may prescribe a medication that is immediately followed by a significant improvement in the patient's symptoms, but the time course of improvement may be too soon for the medication to have had an effect. A major life event also may have led to a significant improvement in the symptom. The patient discussed in Taylor, Chapter 19 of this volume, had multiple problems and showed substantial improvement over the course of therapy, but it was not clear that the therapy led to the improvement. The same

can be said about a doubtful case in which the outcome is not as expected. Here again, several interceding factors can affect the patient in ways that may reduce an otherwise effective approach. Both "clear" and "doubtful" outcomes must be considered as only one source of information about one's practice, but they can serve to help the clinician consider what he or she may be doing well or where he or she might be able to improve.

Peer Review

One approach to evaluating one's performance is to discuss cases with peers routinely during case conferences, during private consultations, or in other settings. The evidential quality of the feedback will be determined by the expertise of the group and the type of information that is presented. Such clinical feedback groups usually focus on particular therapeutic issues or problems rather than on outcomes. Some practice groups may hire individuals with particular expertise, such as an academic psychopharmacologist with a focus on bipolar disorder, to discuss treatment approaches. Such expert advice may be useful, particularly if the expertise is supported by evidence.

Measurement-Based Practice

A more demanding approach is to consider how well a patient is doing relative to some evidence-based standard. Such an evaluation could focus on implementation or outcome. With an implementation-based focus, clinicians would evaluate how well they are adhering to the guidelines and algorithms they favor. With an outcome-based focus, clinicians would focus on how well they are achieving their designated outcomes. An outcome focus has the advantage of helping to ensure that the patients are being provided optimal care, assuming of course that the outcomes are relevant to the patient's needs and that the therapies leading to these outcomes have been shown to be consistent. Outcome-based approaches, with frequent assessments linked to algorithms and other stepped care approaches that delineate and/or suggest practices and decisions, may lead to improved care. This approach has been called *measurement-based care* (Trivedi and Daly 2007; Trivedi et al. 2007). The advantages of measurement-based care are illustrated by Trivedi and Kurian in Chapter 21. This case, written by one of the lead investigators of Sequenced Treatment Alternatives to Relieve Depression (STAR*D) and one of his colleagues, illustrates how a patient's progress is used to determine treatment decisions. Critics of STAR*D might argue that it does not adequately address or incorporate psychosocial issues or interventions, does not deal with comorbidities, and has an important but limited outcome. Even so, patients deserve to achieve *at least* the rate of depression remission and symptom reduction suggested by STAR*D while also achieving gains in other areas. Other cases also indicate that patients' progress can be considered in relation to measures, outcomes, guidelines, and algorithms.

In the cases presented in Parts II and III of this book, we asked the authors to discuss their choice of guidelines and algorithms and reflect on what they learned from the case presented that might affect how, if at all, they could do a better job next time.

Guideline- and Algorithm-Based Psychiatry Practice

With guideline-based psychiatric practice, clinicians would begin by choosing the guidelines or combination of guidelines most relevant to the patient being treated. Clinicians might then "check off" their practices against these guidelines and perhaps at the end of treatment, as illustrated in Chapters 16–19 and 34 of this volume.

For example, suppose a clinician wants to evaluate her care of a patient presenting with both depression and substance abuse. The clinician would identify the appropriate guidelines and outcomes and determine how and when these outcomes would be assessed. In Chapter 8, I discussed how the clinician can identify measures appropriate to his or her practice and patients. For this case, the clinician would focus on depression outcomes and substance use. For instance, the clinician might choose to follow the general American Psychiatric Association depression guidelines but to modify the pharmacology to follow the STAR*D algorithm (see Chapter 21). In accordance with STAR*D, the clinician decides that the patient should experience a 50% reduction in his depressive symptoms by 12 weeks and should be in remission by 6 months, as measured by the Quick Inventory of Depressive Symptomatology self-report version (QIDS-SR) (see Chapter 21 for definitions of remission). The pharmacological intervention would begin with citalopram, with an increase in dose and use of alternative medications

and/or cognitive-behavioral therapy (CBT) as per the STAR*D algorithm (see Chapter 21).

For substance abuse, the clinician decides to generally follow the American Psychiatric Association substance use guidelines but to add a time frame for achieving the predetermined outcomes. Because the guidelines are not very specific, the clinician uses her own standards and experience to define the goals and time frame. For instance, as seen in Table 14–1, for the goal of "motivate the patient to change," the clinician arbitrarily sets a personal standard of achieving this by 3 weeks. The clinician also believes in abstinence and feels that this should be achieved by 2 months into treatment. The clinician agrees that it is important for the patient to "repair disrupted relationships and enhance familial and interpersonal relationships that will support an abstinent lifestyle" and would hope to have the patient do so by 6 months and to help the patient develop social and vocational skills, also by 6 months. For the next goal, "help the patient learn, practice, and internalize changes in attitudes and behaviors conducive to relapse prevention," the clinician wants to accomplish this once the patient has become abstinent before termination. She plans to use a relapse prevention approach.

At week 3, the clinician evaluates how much progress has been made on motivating the patient to take steps toward abstinence. If the patient has not made progress, she reviews how the problem is being approached. The clinician may decide that a different approach is needed (e.g., to put more emphasis on motivational interviewing or establish a new time frame to achieve this goal).

After 8 weeks of therapy, after having generally followed the STAR*D guidelines, she gives the patient another QIDS-SR to assess his depressive symptoms. The QIDS-SR score has declined by more than 50% to a 5, and the depression is in remission. If the patient had not improved, the clinician would continue "down" the STAR*D algorithm with dose escalation, augmentation, and use of CBT, as appropriate and desired by the patient. With this approach, STAR*D achieved an overall remission rate of 67% (Rush et al. 2006).

Following STAR*D makes evaluation of overall practice relatively easy because each case is considered relative to some standards. But many clinicians, for a variety of reasons, might prefer not to use STAR*D or all aspects of STAR*D. Should clinicians using non-STAR*D approaches to treating major depressive disorder be expected to achieve results similar to those with STAR*D (i.e., an overall 67% remission)? In some settings, depending on the patients being treated, this is an unrealistic target. STAR*D indicated that some patients will not benefit even from this very aggressive approach. For instance, the remission rates were 37%, 31%, 14%,

TABLE 14–1. Problem, goals, time frame, and measures

Goal	Time frame[b]	Intervention/activity	Process measure
Motivate the patient to change: patients are in action phase of change, taking steps or abstinent.[a]	By week 3	Motivation interview	Self-report
Achieve abstinence.	At 2 months	Motivation interview/ Cognitive-behavioral therapy	Self-report
Repair disrupted relationships and enhance familial and interpersonal relationships that will support an abstinent lifestyle.[a]	Improvement by 6 months	Family, couples therapy if appropriate	Self-report
Develop social and vocational skills.[a]	Improvement by 6 months	Vocational training	Self-report
Help the patient learn, practice, and internalize changes in attitudes and behavior conducive to relapse prevention.[a]	Before termination	Relapse prevention	Self-report

[a]American Psychiatric Association substance use guidelines.
[b]Time frame added by author.

and 13% for the first, second, third, and fourth acute treatment steps, respectively (Rush et al. 2006). The clinician might take a moment to ask him- or herself: What are my treatment goals for depression, and am I achieving them?

Evaluating one's psychotherapeutic competency is an even more complicated process than evaluating pharmacological practice and outcomes. Suppose a clinician decides that the use of CBT would reduce relapse rates in depressed patients and wants to add this to his or her therapeutic approach. CBT researchers have defined a set of skills appropriate for providing effective CBT and a criterion, representing the sum of these scores, that is arbitrarily considered to reflect competency (Trepka et al. 2004). The instrument is meant to be scored by an experienced rater, but for the practicing clinician, it could be self-scored. Clinicians might consider assessing themselves on how adequately they are applying the basic therapeutic strategies: setting an agenda, providing feedback, being understanding, being interpersonally effective and collaborative, pacing the sessions, and using the time efficiently. Clinicians also might consider how they are teaching skills they wish patients to acquire. For instance, Strachowski et al. (2008) provide an example of how they monitored their patients' use of CBT or engagement in behaviorally activating actions. Assessing practice competence of the more structured interventions, such as CBT, is much simpler than evaluating one's psychotherapeutic competence of less well-defined strategies.

For alcohol abuse, the clinician's first goal (Table 14–1) was to motivate the patient to consider abstinence by week 3. Even though the clinician thought that she did a good job of providing motivational interview techniques, the patient is still drinking and not convinced that he should abstain or even that he has alcoholism. When patients are not progressing as expected, it is reasonable to review diagnosis, assessment, and treatment but also to examine the therapeutic relationship. In this case, the clinician evaluates her treatment according to the Working Alliance Inventory (Horvath and Greenberg 2008). This scale measures client's views of the therapy goals and tasks and of the bond between him or her and the clinician. The clinician finds a large discrepancy between her goals for one aspect of therapy—abstinence from alcohol—and those of the patient—to cut down. She finds that the patient perceives her as caring and understanding. The clinician wonders, however, if she has not been direct enough in confronting the patient's denial about what she perceives as his alcohol dependence.

Overall Practice Performance

If the clinician collected data on a few patients with similar problems and presentations, she could use these types of data to evaluate her overall practice, looking for relevant standards in the scientific literature. However, this can be time-consuming, and such standards may be hard to come by. Although standards for practice seem to be a worthy goal, figuring out how to find and use them is a problem. In the following section, I discuss some approaches, realizing that I am simplifying a very complicated issue and using some relatively easy examples. Yet even these examples illustrate the difficulties of doing evidence-based practice evaluation.

For depression, the clinician might want to assess how the patient is doing compared with STAR*D remission and relapse rates. As noted earlier, an overall target might be 67%. She decides that at least 10 sessions of CBT are necessary to achieve a positive outcome and establishes this as a goal. If fewer than half of her patients are seen for the targeted number of 10 sessions, she might want to examine this aspect of her practice. The clinician has established an abstinence goal at 16 weeks for patients who present with alcohol abuse or dependence and who are drinking. However, she has difficulty finding studies that have used abstinence as a major outcome because many recent alcohol trials report measures such as percentage of days abstinent or days of heavy drinking (Anton et al. 2006). She decides also to use the COMBINE study improvement criteria, which enrolled more than 1,300 patients for a study that compared nine different alcohol abuse treatments (Anton et al. 2006), as one source of information to judge how she is doing. In this study, the main outcomes were percentage of days abstinent over the 8 weeks of the study and time to first heavy drinking day (five standard drinks per day for men; four for women). The mean percentage of days abstinent was about 65%, or 2 days out of 3. The clinician thought that this was a minimal standard, which she defined as the number of patients in her practice with a diagnosis of alcohol abuse or dependence who were sober more than 65% of the time, in the last 2 months of this treatment. She then looked at 10 patients she had treated in the past year

and found that 3 were abstinent for the last 2 months of treatment, and the rest were sober about 50% of the time. On average, the patients were sober more than 65% of the time, meeting the criteria, and had reduced number of heavy drinking episodes. It was difficult for her to know, however, if these results meant that she was doing well or how she could improve her practice.

Checking one's practice against guidelines or algorithms and outcomes is one of many approaches to evidence-based psychiatric practice. The American Psychiatric Association provides a set of tools to make this easier, including a resource for measures (Rush et al. 2008), and the guidelines themselves are designed to improve practice. For illustrative purposes, I translated some of the guidelines into checklists and measured my performance against these guidelines for Chapters 16, 17, and 19.

Performance Measures

Many clinical settings have adopted standards of care and performance measures. These are usually used to determine how well an overall practice of clinicians is doing. For instance, The Joint Commission and the National Association of Psychiatric Health Systems (NAPHS), the National Association of State Mental Health Program Directors (NASMHPD), and the NASMHPD Research Institute, Inc. (NRI), have developed a set of core performance measures for hospital-based inpatient psychiatric services (www.jointcommission.org) that affect all the providers within that system. For instance, inpatient psychiatric health systems need to prove that they screen patients for violence risk, substance use, psychological trauma history, and patient strengths and show that a postdischarge continuing care plan was created and transmitted to the next level of care provider on discharge, among other activities.

Such measures tend to focus on process rather than outcome, although they are presumably developed to ensure better outcomes in the domains assessed. However, they are not necessarily evidence based.

The Department of Veterans Affairs routinely assesses clinician performance on several measures with the goal of 100% compliance to the measure. For instance, through chart review, clinicians may be evaluated on the number of people they screen for alcohol use who were due to be screened or whether they provided a timely suicide assessment and behavior evaluation for patients whose screening results were positive.

The American Psychiatric Association Web site lists several mental health performance measures, as listed in Table 14–2. Links to the source documents and databases are provided at this site.

Standards of care are often related to performance measures. Clinicians also may want to examine how they are doing relative to standards. For instance, the American Diabetes Association, American Psychiatric Association, American Association of Clinical Endocrinologists, and North American Association for the Study of Obesity have created a consensus statement of monitoring weight, blood pressure, glucose, and lipids for patients who are taking second-generation antipsychotics (American Diabetes Association 2004). Many such standards exist relative to mental health practice.

Mistakes and Improvement

Osler urged us to learn from our mistakes, but both identifying mistakes and admitting to them are difficult. Medical errors with serious consequences are common. In 2000, the Institute of Medicine (IOM) reported that 44,000–98,000 people die in hospitals each year as the result of medical errors (Kohn et al. 2000). About 7,000 people per year are estimated to die from medication errors alone—about 16% more deaths than the number attributable to work-related injuries. Unfortunately, very few data exist on the extent of the problem outside of hospitals or in mental health practice settings. The IOM report indicated, however, that many errors are likely to occur outside the hospital. For example, in a recent investigation of pharmacists, the Massachusetts Board of Registration in Pharmacy estimated that 2.4 million prescriptions are filled improperly each year in the state. The IOM emphasized that most of the medical errors are systems related and not attributable to individual negligence or misconduct and that the key to reducing medical errors is to focus on improving the systems of delivering care and not to blame individuals. Many performance measures are designed to reduce errors.

In March 2001, the IOM released a second report on its research into the safety of health care in the United States (Kohn et al. 2001). This report focuses on 13 specific recommendations designed to

TABLE 14–2. Mental health performance measures

Resource	Comments
Quality Measure Inventory. From: Center for Quality Assessment and Improvement in Mental Health (CQAIMH)	Searchable database of more than 300 mental health performance measures
Experience of Care and Health Outcomes Survey (ECHO)	Mental health patient experience-of-care survey instrument
MDD Measures. From: Foundation for Accountability	Legacy documents maintained by the Markle Foundation
Healthcare Effectiveness Data and Information Set (HEDIS) 2008. From: National Committee for Quality Assurance (NCQA)	Health plan–level measures, including several mental health measures
Hospital-Based Inpatient Psychiatric Services (HBIPS) Core Measure Set. From: The Joint Commission	
Health Disparities Collaboratives (HDC) Depression Measures. From: Health Resources and Services Administration (HRSA)	
Behavioral Health Guidelines. From: Institute for Clinical Systems Improvement (ICSI)	MDD and ADHD guidelines, including performance measures
National Quality Measures Clearinghouse (NQMC). From: Agency for Healthcare Research and Quality (AHRQ)	Searchable compendium of performance measures from many domains of health care, including mental health
Physician Consortium for Performance Improvement MDD Measures. From: American Medical Association	
Quality Measures for Schizophrenia. Compiled by New York State Office of Mental Health	
Standards for Bipolar Excellence (STABLE) Project. Maintained by CQAIMH	
VA/DoD Clinical Practice Guidelines. From: U.S. Department of Veterans Affairs	Includes guidelines and measures for MDD, psychoses, and substance use disorder
Washington Circle Substance Use Disorder Measures	Core set of performance measures for alcohol and other drug services for public- and private-sector health plans

Note. ADHD=attention-deficit/hyperactivity disorder; MDD=major depressive disorder.

Source. Adapted from http://www.psych.org/MainMenu/PsychiatricPractice/QualityImprovement/PerformanceMeasures/MentalHealthPerformanceMeasures.aspx.

provide a guide for organizations to use in their efforts to improve patient safety. (An online course is available at www.criticalconceptsusa.com/online_coursesPME/PME_UPDATE.html).

Before clinicians can change practice procedures, they need to be able to identify that an error has been made and admit that one has made or contributed to an error. In this highly litigious practice environment, it is frightening to admit errors, particularly if they had adverse consequences, and to take steps to improve one's practice, which could be seen as an "admission of guilt." Among the barriers to being able to admit errors are resistance to change, fear of discipline, fear of licensure sanction, hindsight bias, need to blame, need to rationalize a negative event, legal system that focuses on fault-finding, and exposure to liability. The issues involved in discussing and disclosing errors with patients, colleagues, and others are beyond the scope of this chapter. The reader is referred to the

IOM report (Kohn et al. 2001) to consider how the IOM recommendations might be used to improve one's practice. The best approach is to prevent errors and for the clinician to monitor practice parameters before having a bad outcome.

However, some minor "mistakes" that clinicians make in practice have no major consequences. In addition to ensuring that they establish practice procedures to avoid serious errors, clinicians also should realize that they could probably do things a little better with almost every patient. The notion that clinicians also should be trying to improve their practice and outcomes is at the heart of evidence-based psychiatric practice. At the end of each case, I asked the authors to reflect on what they might have learned from that case and how they might provide better care to the next patient on the basis of their review. When I compared my actual practice with those suggested by the guidelines I had chosen, I was sometimes surprised to see how many "little" things I had omitted that might have made a difference in the outcome.

References

American Diabetes Association, American Psychiatric Association, American Association of Clinical Endocrinologists, North American Association for the Study of Obesity: Consensus development conference on antipsychotic drugs and obesity and diabetes. Diabetes Care 27:596–601, 2004

Andrews LB, Burruss JW: Core Competencies for Psychiatric Education: Defining, Teaching, and Assessing Resident Competence. Washington, DC, American Psychiatric Publishing, 2004

Anton RF, O'Malley SS, Ciraulo DA, et al: Combined pharmacotherapies and behavioral interventions for alcohol dependence: the COMBINE Study—a randomized controlled trial. JAMA 295:2003–2017, 2006

Cushing H: The Life of William Osler. New York, Oxford University Press, 1940

Horvath AO, Greenberg LS: Working Alliance Inventory (WAI), in Handbook of Psychiatric Measures, 2nd Edition. Edited by Rush AJ Jr, First MB, Blacker D. Washington, DC, American Psychiatric Publishing, 2008, pp 187–191

Kohn LT, Corrigan JM, Donaldson MS (eds): To Err Is Human: Building a Safer Health System. Washington, DC, Institute of Medicine, Committee on Quality of Health Care in America, National Academy Press, 2000

Kohn LT, Corrigan JM, Donaldson MS (eds): Crossing the Quality Chasm: A New System for the 21st Century. Washington, DC, Institute of Medicine, Committee on Quality of Health Care in America, National Academy Press, 2001

Rush AJ, Trivedi MH, Wisniewski SR, et al: Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 163:1905–1917, 2006

Rush AJ Jr, First MB, Blacker D (eds): Handbook of Psychiatric Measures, 2nd Edition. Washington, DC, American Psychiatric Publishing, 2008

Strachowski D, Khaylis A, Conrad A, et al: The effects of directed cognitive behavior therapy on depression in older patients with cardiovascular risk. Depress Anxiety 25:E1–10, 2008

Trepka C, Rees A, Shapiro DA, et al: Clinician competence and outcome of cognitive therapy for depression. Cognit Ther Res 28:143–157, 2004

Trivedi MH, Daly EJ: Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. Drug Alcohol Depend 88 (suppl 2):S61–S71, 2007

Trivedi MH, Rush AJ, Gaynes BN, et al: Maximizing the adequacy of medication treatment in controlled trials and clinical practice: STAR(*)D measurement-based care. Neuropsychopharmacology 32:2479–2489, 2007

This page intentionally left blank

15

Teaching Evidence-Based Medicine and Evidence-Based Psychiatric Practice to Psychiatry Residents

Gregory E. Gray, M.D., Ph.D.
C. Barr Taylor, M.D.

This chapter is intended as a brief overview for residency directors and faculty who are responsible for teaching evidence-based medicine (EBM) to their residents. Other useful sources of information about teaching EBM include the works of Davies (1999), Gray (2001), Green (2000), and Sackett et al. (2000). The Evidence-Based Medicine Resource Center (www.ebmny.org/teach.html) sponsors a Web site that provides access to the Centre for Evidence-Based Medicine teaching materials and presentations, a list of EBM courses on the Web, and other materials. The Accreditation Council for Graduate Medical Education (ACGME) Outcome Project provides a variety of resources related to teaching and assessing EBM and related competencies. The ACGME (2005) also has developed a booklet to promote education in practice-based learning and improvement. The booklet includes a list of resources and examples for teaching and assessing this competency. The American Psychiatric Association also publishes a book providing overviews of the core competencies and how to teach them (Andrews and Burruss 2004).

Why Teach Evidence-Based Medicine and Evidence-Based Psychiatric Practice?

For more than a decade, teaching medical students and residents the fundamentals of clinical epidemiology and EBM has been viewed as a way of enabling them to keep up with the medical literature and of improving clinical care (Evidence-Based Medicine Working Group 1992; Sackett et al. 1991). More recently, however, it has become a required part of residency education in the United States. The ACGME introduced general competencies for residents that are to be included in the requirements for all specialties (Scheiber et al. 2003). The six general areas of core competencies are 1) patient care, 2) medical knowledge, 3) interpersonal and communications skills, 4) practice-based learning and improvement, 5) professionalism, and 6) systems-based practice. As seen in Table 15–1, the core competencies for practice-based learning and improvement for psychiatry are directly related to EBM and evidence-based psychiatric practice (EBPP). For instance, competency 4 states that "the physician shall demonstrate an ability to critically evaluate relevant medical literature," and competency 5 states that the physician shall demonstrate the ability to learn from

TABLE 15–1. Psychiatry core competencies related to evidence-based medicine and evidence-based psychiatric practice

1. The physician shall recognize limitations in his or her own knowledge base and clinical skills and understand and address the need for lifelong learning.
2. The physician shall demonstrate appropriate skills for obtaining and evaluating up-to-date information from scientific and practice literature and other sources to assist in the quality care of patients. This shall include, but not be limited to, the following:
   - Use of medical libraries
   - Use of information technology, including Internet-based searches and literature databases (e.g., MEDLINE)
   - Use of drug information databases
   - Active participation, as appropriate, in educational courses, conferences, and other organized educational activities at both the local and the national levels
3. The physician shall evaluate caseload and practice experience in a systematic manner. This may include the following:
   - Case-based learning
   - Use of best practices through practice guidelines or clinical pathways
   - Review of patient records and outcomes
   - Obtaining evaluations from patients (e.g., outcomes and patient satisfaction)
   - Obtaining appropriate supervision and consultation
   - Maintaining a system for examining errors in practice and initiating improvements to eliminate or reduce errors
4. The physician shall demonstrate an ability to critically evaluate relevant medical literature. This ability may include the following:
   - Use of knowledge of common methodologies employed in psychiatric and neurological research
   - Researching and summarizing a particular problem that derives from the physician's caseload
5. The physician shall demonstrate the ability to do the following:
   - Review and critically assess scientific literature to determine how quality of care can be improved in relation to practice (e.g., reliable and valid assessment techniques, treatment approaches with established effectiveness, practice parameter adherence). Within this aim, the physician shall be able to assess the generalizability or applicability of research findings to patients in relation to their sociodemographic and clinical characteristics.
   - Develop and pursue effective remediation strategies that are based on critical review of scientific literature
   - Learn from one's own and other specialties

Source. Adapted from Scheiber et al. 2003.

one's practice, the subject of the previous chapter. In addition, one of the systems-based practice requirements is to be able to use practice guidelines.

What to Teach

The ACGME general competencies require that psychiatry residents become competent in using the EBM process to answer questions about therapies and diagnostic tests. At a minimum, therefore, residents should be taught how to formulate a clinical question (Chapter 3); perform a literature search (Chapter 4); appraise clinical trials, systematic reviews, guidelines, and diagnostic tests (Chapters 5–7 and 9); and apply the results to their patients. A more complete course would include appraisal of articles on disease frequency (Chapter 10), etiology or harm (Chapter 11), and prognosis (Chapter 12) and learn-

ing how to apply these results to their practice (Chapter 13 and as illustrated in the cases in the remainder of this volume).

In providing this instruction, it is important to realize that most residents are not interested in clinical epidemiology or in learning some of the more advanced EBM skills (Guyatt et al. 2000). The goal instead should be to ensure that residents are able to find answers efficiently by using the preappraised information sources we described in Chapter 4 (Evans 2001; Guyatt et al. 2000) and to understand how to use measures, algorithms, guidelines, and similar resources in practice. However, at times, preappraised sources will not yield an answer, or it may be necessary to discuss the evidence in more detail with a colleague or patient. Under these circumstances, it is necessary to have the kind of deeper knowledge that is covered in an advanced course (Woodcock et al. 2002). Furthermore, a deeper understanding of EBM allows clinicians to better appraise the quality of guidelines developed by others and to better understand and incorporate the results from systematic reviews into clinical practice (Doust and Silagy 2000; Laupacis 2001).

One of the more important aspects of teaching EBM to psychiatry residents is teaching the philosophy of EBM and EBPP, not just the methodology. Residents often have some misconceptions about EBM (Bilsker and Goldner 1999; Padrino 2002). In addition to providing an overview of what EBM is and what it is not (see Taylor and Gray, Chapter 1 in this volume) and discussing some of the misconceptions about EBM (Straus and McAlister 2000), it may be helpful to distribute the humorous article by Isaacs and Fitzgerald (1999), in which they describe alternatives to EBM, such as *eminence-based medicine* ("experience is worth any amount of evidence"), *vehemence-based medicine* ("substitution of volume for evidence"), and *eloquence-based medicine*. This article is also helpful to residents in coping with situations in which the opinions of their supervisors are at odds with the evidence in the literature.

Methods of Teaching Evidence-Based Medicine and Evidence-Based Psychiatric Practice

The ACGME Outcome Project has a Web site that showcases examples of activities that may be used to teach and assess the general competencies (www.acgme.org/outcome/implement/rsvp.asp). The descriptions are provided directly by the program or institution. The user is advised to communicate directly with the contact person listed for each project for further information. Most of the examples are not related to practice-based learning and do not directly address psychiatry or psychiatric competencies.

Seminars and Small-Group Sessions

Some material can best be introduced in seminars or in small-group sessions. In the University of Southern California program, the topics (e.g., critical appraisal of clinical trials) were presented to a small group in the form of a brief presentation and discussion. This was then followed by the critical appraisal of an article related to the topic under discussion (e.g., an article reporting the results of a randomized controlled trial). To do the appraisal, the residents were given a copy of the article along with a worksheet that incorporates the appropriate critical appraisal guidelines. After the residents had completed the worksheet, individual items on it were discussed. Although such sessions are useful for introducing a topic or method, they do not provide sufficient practice for the residents to become proficient; for this, other approaches are more useful.

Computer Laboratory and Library Sessions

The best way to become familiar with the various databases and how to use them is through hands-on practice, not by reading about them or through a lecture. Most medical schools have computer laboratories or library classrooms equipped with computers that can be used for demonstrations and hands-on practice in a group setting. Most medical libraries offer classes on searching MEDLINE and other databases, and librarians will often customize courses for the needs of particular users.

Evidence-Based Journal Clubs

Journal clubs, a staple of residency education programs, are the most common sites for teaching critical appraisal skills (Green 1999, 2000). Sackett et al. (2000) have described a novel approach to journal clubs—the evidence-based journal club—as a method for teaching the EBM process.

In an evidence-based journal club, the starting point is a clinical question suggested by one of the

residents, preferably one that is based on an actual patient care issue. The group agrees on the 4-part PICO question (a question that uses the mnemonic aid "patient/problem, intervention, comparison, and outcome"), and a resident is assigned to conduct a literature search to find the best evidence to answer the question. The resident assigned to the task tries to find the evidence that is highest in the hierarchy (Table 4–1, Chapter 4). At the next journal club meeting, the resident distributes the article that was found, and the group as a whole uses a critical appraisal worksheet to evaluate the article. The question, search process, results, and applicability are then discussed.

The evidence-based journal club format has become popular in psychiatry training programs in the United Kingdom (Dhar 2001; Walker 2001; Warner and King 1997).

Rounds and Supervision

It is generally recommended that EBM not be taught in isolation but instead be incorporated in other clinical teaching activities (Dobbie et al. 2000; Green 1999, 2000; Sackett et al. 2000). In some institutions, this can be done through real-time literature searches during rounds or supervision, but on-line database access is still the exception in most hospital settings (Green 2000). Sackett et al. (2000) and the Evidence-Based Medicine Working Group (1992), for example, describe settings in which on-line searches can be conducted during rounds. Chapter 33 in this volume provides an example of EBM on a consultation-liaison service.

The Department of Psychiatry at Duke University has a model for providing the weekly Chairman's Rounds in an evidence-based format (http://psychiatry.mc.duke.edu/Residents/Quest.html).

The issues for teaching EBPP are more complicated than are those for teaching EBM, and few curricula exist on how to do this. Dr. Tan, who was supervised by the author, provides an example of how a resident-in-training can incorporate aspects of EBPP to develop a treatment plan, determine treatment strategies, and evaluate outcome (see Chapter 34).

Online Teaching Resources

A variety of Web sites provide useful information (such as curriculum guides and online syllabi) for the teaching of EBM (Table 15–2).

An Example of Teaching Evidence-Based Medicine on an Inpatient Service

Mascola (2008) described a curriculum integrated into an inpatient teaching service. The trainees are e-mailed a syllabus before the rotation that specifies the EBM objectives to be learned, practiced, and assessed during the 1-month rotation. The syllabus includes six brief discussion modules designed to teach residents how to apply the key elements of the EBM model in a busy clinical setting. EBM knowledge and skills are assessed before and after the rotation. In a pilot study, Mascola (2008) reported that the knowledge and skills increased significantly relative to baseline in a sample of 24 consecutive trainees who underwent the program. Chapter 33 illustrates how Mascola used EBM on a consultation-liaison service.

A Model Curriculum

The Duke University Department of Psychiatry identifies the teaching of EBM as one of its three primary missions. According to the Web site (www.psychres.duke.edu/residency/general.html):

> Interns begin with an EBM Start-Up course in July of their PGY-1 year, followed by intensive practice writing CATs or Critically Appraised Topics for presentation in our weekly Chairman's Rounds. Our chief residents receive training in how to teach EBM and meet with the presenting residents to guide him/her through the process of selecting a relevant clinical problem from their patient experience, forming a researchable question, searching the literature for the best evidence and appraising that evidence for its validity. Residents practice generating the most relevant summary statistic for the study's results and evaluating the applicability of the study to their patient. Residents receive additional practice writing CATs for other seminars and conferences throughout their training. And an EBM Review course for all residents runs year-round during our Academic Half-day.…[R]esidents are challenged to practice these skills on the wards and in the clinic. EBM-trained attendings and supervisors apply an evidence-based approach to resident supervision, whether the topic is a medication or a psychosocial therapy question.

Aspects of this program are described in March et al. (2007). The grand rounds are presented as evidence-based summaries and reviews and posted on-line as noted earlier.

TABLE 15–2. Useful sites for teachers of evidence-based medicine

Organization	Address
Centre for Evidence-Based Medicine (Oxford)	http://www.cebm.net/index.asp
Centre for Evidence-Based Medicine (Toronto)	http://www.cebm.utoronto.ca/teach
Centre for Evidence-Based Mental Health	http://www.cebmh.com
Evidence-Based Medicine Resource Center	http://www.ebmny.org

Evaluating Evidence-Based Medicine Skills

Most evaluations of EBM have focused on critical appraisal skills of knowledge of EBM terminology and concepts, generally using multiple-choice examinations (Green 1999, 2000; Hatala and Guyatt 2002). Such approaches are subject to many criticisms, including the lack of validation of instruments and failure to assess the entire EBM process.

One promising approach for appraising a broader range of EBM skills is that of Smith et al. (2000), who developed a written test that focuses on four different skills: formulation of questions, literature searches, quantitative understanding, and appraisal of study quality and clinical relevance. Other authors have suggested the use of objective structured clinical examination–type stations to assess specific skills (Dobbie et al. 2000).

Evaluating whether residents are applying EBM skills in their day-to-day practices is as yet an unmet need. Such assessments generally have relied on self-reports, which may not be accurate (Green 1999, 2000). For example, Cabell et al. (2001) found that resident self-reports overestimated the actual use of online databases.

The ACGME Outcome Project (Accreditation Council for Graduate Medical Education 2001) provides several approaches to assessing practice-based learning and improvement competency. Included are self-administered diaries, portfolios, and instruments to assess opinions about EBM and information technology; and knowledge of EBM concepts and critical appraisal skills.

Evaluating Evidence-Based Medicine Teaching Programs

The Centre for Evidence-Based Medicine has a freely accessible Web site (www.cebm.utoronto.ca) developed by the Department of Medicine at Toronto General Hospital. The goal of the Web site is "to help develop, disseminate, and evaluate resources that can be used to practise and teach EBM for undergraduate, postgraduate and continuing education for health care professionals from a variety of clinical disciplines." A self-evaluation toolkit that can be used by the teachers of EBM is available on the Web site. It suggests that EBM educators ask themselves questions such as the following:

- When did I last issue an educational prescription?
- Am I helping my trainees learn how to ask answerable questions?
- Am I teaching and modeling searching skills?
- Am I teaching and modeling critical appraisal skills?
- Am I teaching and modeling the generation of critically appraised topics?
- Am I teaching and modeling the integration of best evidence with my clinical expertise and my patients' preferences?
- Am I developing new ways of evaluating the effectiveness of my teaching?
- Am I developing new EBM educational material?
- Am I a member of an EBM-style journal club?
- Have I participated in or tutored at one of the workshops on how to practice or teach EBM?
- Have I joined the evidence-based health e-mail discussion group?
- Have I established links with other practitioners or teachers of EBM?

The ACGME Web site offers examples of assessment tools, as noted earlier.

Resident Views of Evidence-Based Medicine

Psychiatry residents often have some exposure to EBM as medical students, although this is not neces-

sarily true for international medical graduates. Residents often have some ambivalence toward EBM, wondering if it ignores the humanistic side of psychiatric practice (Bilsker and Goldner 1999). However, as they develop a better understanding of EBM and the roles of clinical judgment and patient preference, coming to realize that EBM and patient-centered care are complementary, such concerns generally lessen (Bilsker and Goldner 1999; Hope 2002).

Some residents are intimidated by the quantitative emphasis of EBM (Bilsker and Goldner 1999). This generally can be overcome by setting realistic goals and by remembering that the focus should be on preparing users of evidence, not researchers (Guyatt et al. 2000; Sackett et al. 2000).

A barrier to teaching EBM and EBPP may be the negative opinions of some faculty members (Ball 1999; Evidence-Based Medicine Working Group 1992). Faculty development efforts may be needed to address the negative opinions (Bilsker and Goldner 1999; Evidence-Based Medicine Working Group 1992; Neale et al. 1999), although some residents are able to educate their supervisors about the concepts (Ball 1999).

Finally, it is important to emphasize that many residents find the practice of EBM to be empowering (Ball 1999; Padrino 2002). As Padrino (2002) said, "I feel more confident in my medical decisions when I can say 'the data show this' or 'the data show that.' Even when I have to say 'there are no data for this,' I feel my decision is more valid" (p. 13). Being able to incorporate the best evidence from the research literature into patient care decisions is what EBM is all about, and residents are quite capable of learning how to use evidence to improve their clinical decision making.

References

Accreditation Council for Graduate Medical Education: Practice-based learning and improvement: assessment approaches [ACGME Outcome Project Web site]. 2001. Available at: http://www.acgme.org/outcome/assess/PBLI_Index.asp. Accessed October 29, 2002

Accreditation Council for Graduate Medical Education: Advancing Education in Practice-Based Learning & Improvement: An Educational Resource From the ACGME Outcome Project. 2005. Available at: http://www.acgme.org/outcome/implement/complete_PBLIBooklet.pdf. Accessed June 28, 2009

Andrews LB, Burruss JW: Core Competencies for Psychiatric Education: Defining, Teaching, and Assessing Resident Competence. Washington, DC, American Psychiatric Publishing, 2004

Ball C: Evidence-based medicine on the wards: report from an evidence-based minion. Evid Based Med 3:101–103, 1999

Bilsker D, Goldner EM: Teaching evidence-based practice in mental health. Evid Based Ment Health 2:68–69, 1999

Cabell CH, Schardt C, Sanders L, et al: Resident utilization of information technology: a randomized trial of clinical question formation. J Gen Intern Med 16:838–844, 2001

Davies P: Teaching evidence-based health care, in Evidence-Based Practice: A Primer for Health Care Professionals. Edited by Dawes M, Davies P, Gray A, et al. New York, Churchill Livingstone, 1999, pp 223–242

Dhar R: Evidence-based journal clubs and the critical review paper: candidate's perspective. Psychiatr Bull 25:67–68, 2001

Dobbie AE, Schneider ED, Anderson AD, et al: What evidence supports teaching evidence-based medicine? Acad Med 75:1184–1185, 2000

Doust JA, Silagy CA: Applying the results of a systematic review in general practice. Med J Aust 172:153–156, 2000

Evans M: Creating knowledge management skills in primary care residents: a description of a new pathway to evidence-based practice. ACP J Club 135:A11–A12, 2001

Evidence-Based Medicine Working Group: Evidence-based medicine: a new approach to teaching the practice of medicine. JAMA 268:2420–2425, 1992

Gray JAM: Evidence-Based Healthcare: How to Make Health Policy and Management Decisions, 2nd Edition. New York, Churchill Livingstone, 2001

Green ML: Graduate medical education training in clinical epidemiology, critical appraisal, and evidence-based medicine: a critical review of curricula. Acad Med 74:686–694, 1999

Green ML: Evidence-based medicine training in graduate medical education: past, present, and future. J Eval Clin Pract 6:121–138, 2000

Guyatt GH, Meade MO, Jaeschke RZ, et al: Practitioners of evidence-based care. BMJ 320:954–955, 2000

Hatala R, Guyatt G: Evaluating the teaching of evidence-based medicine. JAMA 288:1110–1112, 2002

Hope T: Evidence-based patient choice and psychiatry. Evid Based Ment Health 5:100–101, 2002

Isaacs D, Fitzgerald D: Seven alternatives to evidence based medicine. BMJ 319:1618, 1999

Laupacis A: The future of evidence-based medicine. Can J Clin Pharmacol 8 (suppl A):6A–9A, 2001

March JS, Chrisman A, Breland-Noble A, et al: Using and teaching evidence-based medicine: the Duke University child and adolescent psychiatry model. Child Adolesc Psychiatr Clin N Am 14:273–296, viii–ix, 2005

March JS, Szatmari P, Bukstein O, et al: AACAP 2005 Research Forum: speeding the adoption of evidence-based practice in pediatric psychiatry. J Am Acad Child Adolesc Psychiatry 46:1098–1110, 2007

Mascola AJ: Guided mentorship in evidence-based medicine for psychiatry: a pilot cohort study supporting a promising method of real-time clinical instruction. Acad Psychiatry 32:475–483, 2008

Neale V, Roth LM, Schwartz KL: Faculty development using evidence-based medicine as an organizing curricular theme. Acad Med 74:611, 1999

Padrino S: Evidence-based psychiatry: fad or fundamental? Psychiatr News 37(16):13, 2002

Sackett DL, Haynes RB, Guyatt GH, et al: Clinical Epidemiology: A Basic Science for Clinical Medicine, 2nd Edition. Boston, MA, Little, Brown, 1991

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

Scheiber JC, Kramer TAM, Adamowski SE: The implications of core competencies for psychiatric education and practice in the US. Can J Psychiatry 48:215–221, 2003

Smith CA, Ganschow PS, Reilly BM: Teaching residents evidence-based medicine skills: a controlled trial of effectiveness and assessment of durability. J Gen Intern Med 15:710–715, 2000

Straus SE, McAlister FA: Evidence-based medicine: a commentary on common criticisms. CMAJ 163:837–841, 2000

Walker N: Evidence-based journal clubs and the critical review paper. Psychiatr Bull 25:237, 2001

Warner JP, King M: Evidence-based medicine and the journal club: a cross-sectional survey of participants' views. Psychiatr Bull 21:532–534, 1997

Woodcock JD, Greenley S, Barton S: Doctors' knowledge about evidence-based medicine terminology. BMJ 324:929–930, 2002

This page intentionally left blank

II

Case Studies: Evidence-Based Medicine Applied to Major DSM-IV Disorders

This page intentionally left blank

16

Introduction to the Cases

C. Barr Taylor, M.D.

The cases presented in the remainder of this volume are designed to illustrate the principles and practices of evidence-based psychiatric practice (EBPP). As discussed in Chapter 1, EBPP is based on the following:

1. Accurate evaluation and treatment planning
2. Consideration of treatment algorithms, guidelines, and best practices when planning and providing treatment
3. Use of measures to determine progress of an individual patient
4. Review of a patient's progress against personal and published standards
5. Consideration of "evidence and experts" in clinical decisions
6. Expertise in providing a range of therapies or in making them available through other resources or providers

In addition to improving patient care, EBPP is a method of achieving practice-based learning and improvement, one of the core competencies now expected of all graduating residents. Having directed a residency training program for over a decade, I have found that this is one of the more difficult competencies for faculty to teach and practitioners to enact. I began this project in part to find better ways to teach residents how to learn from patients and improve their care and found that it helped improve my own practice. These cases are presented to illustrate how I and others have used EBPP in our practice.

I asked each of the authors to follow a general outline as follows:

Setting. Provide a brief description of your practice setting, patient population/focus, and treatment orientation/philosophy.

Illustration. List several applications of expert/evidence-based psychiatry that will be illustrated.

Chief complaint. Use present tense for statement of the chief complaint.

Present illness. The case should be "real" but, for ethical reasons, the identity of the individuals should be changed so that they cannot be identified by others (e.g., change age, sex, occupation, personal details). While real data are ideal, it may be necessary to combine data from patients. Keep it brief.

Other significant findings from assessment. Present only relevant findings (including negative ones if they are pertinent). Keep it brief.

Differential diagnosis (if appropriate).

DSM-IV-TR assessment.

Treatment plan considerations. Identify the main and secondary problems and issues. As appropriate for your practice and the patient, describe the guidelines, algorithms, measures, or other sources of EBPP you used.

Treatment goals, measures, and methods. Include a time frame for measurement/improvement, preferably as a table.

Course. Provide a brief discussion of the course. An important aspect of these cases is to let the reader into how you were thinking about the problem over the course of therapy. If you deviated from expert practices, why? How did you use the manual, self-help materials, brochures, and other sources, if at all? How did you consider problems that emerged and were not covered by guidelines and algorithms? How did you coordinate care if you referred to other practitioners (e.g., for marital therapy or substance abuse treatment)? Were there issues you researched using EBM [evidence-based medicine] methods?

Figure/graph/table of improvement (if appropriate and available).

Summary (including use of guidelines/algorithms, other issues).

Ways to improve practice. What did you learn from this case (if anything) that might help you improve your practice?

One challenge of writing about real clinical cases, as mentioned earlier, is the need to protect the privacy of the individuals being discussed. For this reason, I have asked the authors to change the age, sex, job, revealing personal details, and, even when it was not essential, the history of the patient and the course of treatment.

Another challenge is to make the case "come alive." In my practice I usually cover the core of the American Psychiatric Association (2002) practice guidelines baseline assessment and obtain a detailed history and mental status. However, in my cases, I present only information relevant to the case or illustrative of the issues addressed. I asked the other authors to do the same. I also do not discuss general therapy issues and techniques, the focus of much of my work with patients, and I didn't expect the other authors to. To do so would make the cases very long and cover material familiar to most clinicians.

When I first began working on this book, my initial idea was to write up a dozen or so illustrative cases taken from my own practice. In the end, I decided to restrict the cases I present to anxiety, eating, and affective disorders and comorbidities, the most common problems presenting in my clinic, and to invite specialists in other areas to prepare evidence-based case reports around problems they routinely address so that examples of evidence-based treatment would be available for many of the major DSM-IV-TR (American Psychiatric Association 2000) disorders where guidelines are available. I also felt it was important to provide examples of EBPP in a variety of different treatment settings.

The cases address a number of ways to undertake EBPP. Table 16–1 presents an overview of the case diagnoses, the guidelines/algorithms followed, and the outcome measures. I encouraged the authors to present real cases, including those where the outcome was not as expected. In such cases, I also asked the authors to consider what they might have done differently, if anything. The cases are meant to be illustrative of how to treat the cases presented. I suspect the readers will have their own views as to how they might have approached the cases in different and perhaps better ways.

References

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

American Psychiatric Association: Practice Guidelines for the Treatment of Psychiatric Disorders, Compendium. Washington, DC, American Psychiatric Association, 2006

TABLE 16–1. Case authors, disorders, guidelines/algorithms, and outcome measures

Chapter	Authors	Disorder(s)/Issues	Guidelines/Algorithms	Outcome measures
17	Taylor	Panic disorder	NICE—anxiety disorders	$QIDS\text{-}SR_{16}$ Self-report: panic attack frequency, dizziness, depression (severity)
18	Taylor	MDD Bulimia Relationship issues	APA—depression, STAR*D NICE—eating disorders	$QIDS\text{-}SR_{16}$ Self-report: bulimic episodes, relationship issues (satisfaction), depression
19	Taylor	MDE Weight loss/ food refusal Relationship issues	APA—depression	$QIDS\text{-}SR_{16}$ Weight Self-report: difficulty swallowing, depression
20	Taylor	Anorexia OCD Depression Psychosis NOS	APA—anorexia APA—OCD	$QIDS\text{-}SR_{16}$, Weight Obsessions (frequency/severity) BPRS Depression
21	Trivedi & Kurian	MDD	STAR*D	$QIDS\text{-}C_{16}$ FIBSER
22	Rothschild	MDD	APA—MDD	BPRS DAS Ham-D Stress, self-report
23	Wang & Ketter	Bipolar II	TIMA	Bipolarity Index, self-report: mood and elation
24	Das & Koran	OCD MDD	APA—OCD	Y-BOCS Ham-D
25	Rait & Glick	Schizophrenia	APA—schizophrenia (section on family therapy)	Global functioning Global relationship functioning
26	Lembke & Humphreys	Substance abuse	VHA-DoD—substance use disorders	Abstinence
27	Reich	Personality disorders	APA—borderline personality disorder APA—OCD WFSBP—personality disorders	Self-report: personality disorder symptoms
28	Marino & Mitchell	Bulimia	APA—bulimia	Binge/purge frequency

TABLE 16–1. Case authors, disorders, guidelines/algorithms, and outcome measures

Chapter	Authors	Disorder(s)/Issues	Guidelines/Algorithms	Outcome measures
29	Lindner & Lindley	PTSD Substance abuse	VHA-DoD—PTSD	PTSD Checklist Alcohol abstinence Hours of sleep Housed Employed
30	Tannenbaum & Spiranovic	Depression Panic disorder	APA—depression APA—panic disorder RANZCP—depression NICE—depression	Depression Severity Questionnaire Quality of Life Side Effects Scale
31	May & Reynolds-May	Postpartum depression	ACOG Various others	Relationship satisfaction
32	Shanteau	Bipolar I	TMAP	PHQ-9 Self-report
33	Mascola	Dementia, delirium, or depression	Various evidence-based practices on a C-L service	Structured clinical interview
34	Tan	Alcohol abuse, depression	APA—substance use disorders	Alcohol use, alcohol craving, depression

Note. ACOG=American College of Obstetricians and Gynecologists; APA=American Psychiatric Association; BPRS=Brief Psychiatric Rating Scale; C-L=consultation-liaison; DAS=Delusional Assessment Scale; FIBSER=Frequency, Intensity, and Burden of Side Effects Rating; Ham-D=Hamilton Rating Scale for Depression; MDD=major depressive disorder; MDE=major depressive episode; NICE=National Institute for Health and Clinical Excellence; NOS=not otherwise specified; OCD=obsessive-compulsive disorder; PHQ-9=Patient Health Questionnaire (nine questions); PTSD=posttraumatic stress disorder; QIDS-C_{16}=Quick Inventory of Depressive Symptomatology–Clinician Rating; QIDS-SR_{16}=Quick Inventory of Depressive Symptomatology–Self-Report; RANZCP=Royal Australian and New Zealand College of Psychiatrists; STAR*D=Sequenced Treatment Alternatives to Relieve Depression; TIMA=Texas Implementation of Medication Algorithms; VHA-DoD=Veterans Health Administration, Department of Defense; WFSBP=World Federation of Societies of Biological Psychiatry; Y-BOCS=Yale-Brown Obsessive Compulsive Scale.

17

Panic Disorder

C. Barr Taylor, M.D.

Introduction

General Therapeutic Approach

I am an "eclectic psychiatrist," expert in a number of types of therapy but with a bias toward more empirically derived approaches. In general, I prefer to consider a set of techniques that may be helpful for a particular patient rather than to rigidly follow any specific school or approach. I usually provide both psychotherapy and psychopharmacology.

I generally see patients for 50 minutes. I begin my session with a review of how the patient has being doing since the last visit, noting any important events and issues. If the patient has completed the presession assessment in the waiting room, I quickly go over it; otherwise, I review salient aspects (such as depression and/or anxiety and target symptoms or behaviors), medication (side effects, adherence, change in other medications), and homework (if assigned). I then set an agenda, beginning with therapeutic issues related to the most pressing issues raised by the patient, but ensure that I also have time for teaching and practicing new therapeutic strategies, establishing goals and assignments, and setting up a new appointment. I look for indications that the therapeutic alliance is strong. At times the patient may be in crisis, and I drop this agenda to spend as much time as needed to deal with urgent issues or other therapeutic issues that may arise.

Charting Patient Progress

In preparing cases for this book, I decided to keep graphs or charts of the patient's progress and to combine three types of measurements:

1. *Simple scales of depression and anxiety.* As appropriate, I ask the patient, if he or she has not completed the form in the waiting room: "In the past week, what was your average level of depression, where 0=none/no symptoms to 10=severe?" I use the same question adapted for anxiety. (If you use this system it is important to "anchor" the points by asking the patient what his or her point of reference is in rating anxiety or depression as moderate or severe.) I also ask the patient what his or her maximum anxiety or depression was in the past week if appropriate, but I don't graph the data.

2. *Standard measures.* If the patient's problems are appropriate for the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) algorithm, I have the patient complete, or I complete with the patient, the Quick Inventory of Depressive Symptomatology–Self-Report ($QIDS\text{-}SR_{16}$) every 6 weeks or so. I may add another standard measure to assess a problem, such as the Yale-Brown Obsessive Compulsive Scale for obsessive-compulsive disorder.

3. *Measures of target symptoms/problems.* I use one or two other measures targeted specifically to the patient's problems. I may monitor the frequency of a symptom, asking the patient, "In the last week, how often did you experience this?" and have the patient rate it as "never," "almost never," "a few times a week," or "a few times a day." I may also plot actual frequencies of behaviors or symptoms, such as panic attacks or binges. Or I may monitor the severity of a symptom by asking, "On average, how severe (bothersome) was the symptom over the past week?"

I have the patient rate it from 0=not bad at all to 10=severe. Working with the patient, I may develop some other simple measure to reflect the patient's progress toward his or her goals. For instance, if I am dealing with a patient with social phobia who wants to date more, I may decide to use a general measure of how comfortable he feels in dating. I may also use a measure that focuses on satisfaction, for instance, "How satisfied are you with your relationships with women you date?" The data will give me, and the patient, an idea of how he is doing and help determine if I need to change treatment strategies, for example, if the patient is not adequately improving relative to his or my expectation of what progress would be appropriate at that point in therapy. I provide examples of all these types of measures collected over time in the following cases.

There are many challenges in collecting such data, and I believe we should use the data we collect. One challenge is to relate the patient's visit number to the week he or she has been in therapy. To make it easier, I attach a simple chart to the front of the case that looks like this:

Visit	Date	Week
1	09/07/06	0 (baseline)
2	09/13/06	1
3	10/05/06	6
4	10/18/06	13
(etc.)		

The chart allows me to plot the data, by visit, in real time. I usually share the information on the chart with the patient. Doing so allows me to review the patient's goals and keeps me focused on helping the patient achieve progress.

In truth, keeping up even this simple data system is a bit of a nuisance. Patients may resent filling out the forms or find them repetitive or unnecessary, a session may run over, or I may have limited time between sessions or other more urgent issues to attend to. My first priority in charting is to complete the clinical note; the graph is optional. However, it is rarely necessary to keep data at each patient session.

Guidelines

Having been trying to focus on evidence-based psychiatric practice for the past several years, I thought I was familiar with the guidelines. However, in formally reviewing my cases against guidelines, I find that I sometimes deviate more than I should. I felt the reader would benefit from my discussion about these deviations. I decided to turn the guidelines into checklists and then to examine my practice against the guideline as a standard. I used this review as one source for evaluating how I could do a better job with a similar patient in the future.

When patients are not improving as I would expect them to based on my experience or some other source of information, I ask myself: Is my diagnosis/assessment correct? Are my therapeutic choices appropriate and/or being applied (e.g., are there adherence or implementation problems)? Are there problems in the therapeutic relationship? Searches applying methods of evidence-based medicine can be useful to help me determine if I have missed some useful information or approach that might aid the patient.

Setting

I practice in a general outpatient clinic in a private university medical center department of psychiatry. The clinic serves mostly insurance patients. I have set up my office so that my computer is to my left and my patients sit to my right in a chair at the corner of my desk.

Illustration

This study illustrates:

- Application of guidelines to a single disorder
- Use of manualized treatments/self-help
- Monitoring of symptoms
- Dealing with relapse

Chief Complaint

Mr. C.W. is a 32-year-old male engineer who presents with a chief complaint of panic attacks.

Present Illness

Mr. C.W. said that his problems had begun about 3 months prior to presentation when he was driving to work and crossing over a bridge. He experienced a sudden, intense feeling of panic that was accompanied by a sharp pain in the back of his head and a rapid heart rate. He drove himself to an emergency room (ER) where he received medical tests and "no abnormalities" were found. A week later he had another attack when he was again driving to work.

He returned to the ER and received a CT scan, chest X ray, ECG, and other tests, which all yielded "normal" results. He began to find it impossible to cross over the bridge on the way to work without being accompanied by a friend, and he started to avoid work altogether, even though he had not had another panic attack. He also began to experience severe dizziness nearly constantly. He went to his internist, who told him he had anxiety and prescribed citalopram, starting at 10 mg/day, and clonazepam 0.5 mg prn. On the day after the patient took the citalopram he became "severely" depressed, felt like crying, and had suicidal thoughts. He stopped the citalopram on his own but stayed on the clonazepam. He continued to be driven to work by a friend and had a number of episodes when he felt his heart was beating fast on his way to work even when not on a bridge. He then self-referred to the psychiatry outpatient clinic, where I first saw him.

The patient had been keeping detailed notes since the first event. For instance, his notes from one day were as follows:

- 6:45 A.M. heart beating fast
- 7:30 A.M. anxiety, dizzy
- 8:00 A.M. at work, dizzy
- 12:00 P.M. weak and faint
- 1:00 P.M. weak and faint
- 7:00 P.M. feel better, dizzy

When asked what he thought was causing the problem, he suggested, "Work stress, maybe chemicals at work, or maybe I have a brain tumor."

Other Significant Findings From Assessment

The patient exercised one to two times per week (jogging for about 40 minutes on Saturday and Sunday). He was happy with his diet and weight. He reported being isolated, although he socialized with people at work on Friday nights and sometimes on the weekend. He considered himself to be an agnostic. He denied smoking, drinking, or taking any over-the-counter or herbal or other nonprescribed products or medicines. He avoided any foods or drinks with caffeine. There were no medical problems contributing to onset of panic attacks. He reported moderate and increasing stress at work, with long hours. His STAR*D QID-SR_{16} score = 8 (low to moderate depression; see Chapter 21). On his baseline form, he rated his average level of depression over the past week as being about a 3 (moderate, on a scale of 0 = no symptoms and 10 = severe symptoms) and his maximum depression about the same. He rated his average anxiety as being 5/10 with his maximum as 8/10. He said he experienced panic a few times a week. He rated his dizziness as a 6/10.

DSM-IV-TR Diagnosis

Axis I	Panic disorder with agoraphobia
Axis II	None
Axis III	None (cause of dizziness was not clear)
Axis IV	Work stress; social isolation
Axis V	Global Assessment of Functioning (GAF) score: 50

Treatment Plan Considerations

Panic Attacks/Avoidance

On first assessment Mr. C.W. seemed to present a fairly straightforward case of panic disorder with agoraphobia and secondary depression. He was having about two to three "panic episodes" every week although these were rarely full-blown panic attacks, perhaps because of his avoidance and use of benzodiazepines.

Selecting a Guideline

Having reviewed both the American Psychiatric Association (APA) and National Institute for Health and Clinical Excellence (NICE) guidelines (and others) for an article for *BMJ* (Taylor 2006), I ended up preferring the NICE guideline, which was most consistent with my own practice and publications. I provide the guideline at the end of the case and review where I adhered to or deviated from the guideline (www.nice.org.uk/Guidance/CG22).

Given that the guideline lists cognitive-behavioral therapy (CBT) as the therapy of first choice, the first major therapeutic decision for the patient was to determine if medications were appropriate for this case and, if so, which medications to use. CBT would certainly be appropriate in the long run, but would it be sufficient to deal with the patient's distress? The patient was quite impaired (GAF = 50), requiring a friend to drive him to work and feeling miserable most of the day. He also had a number of depressive symptoms (feelings of sadness, fatigue, being slowed

down, and thoughts that life is empty or not worth living). I felt he would benefit from an adequate trial of antidepressant both for the panic and for the depressive symptoms. The patient also was reluctant to use only CBT, and his job required some flexibility as to when I could schedule sessions. The guideline does not suggest that CBT be used in lieu of medication but suggests that CBT should be used first.

Having decided to use a medication, should I choose a selective serotonin reuptake inhibitor (SSRI) other than citalopram or start on another type of antidepressant? The patient had become very depressed on the citalopram when he first tried it. (I made a mental note to see what I could find out about using another SSRI when a patient had a "depressive reaction.") The patient would not try citalopram again, even at half the dose, so I decided to start him on sertraline at a very low dose (a quarter of a 25 mg tablet for 3 days followed by half a tablet for 3 days). The patient was told to stop the medication if he felt worse.

From the NICE guideline perspective, the use of the clonazepam was problematic. The guideline states that benzodiazepines are not to be used, on the basis that "outcome studies argue that patients given benzodiazepines do worse over time." The guideline would imply that I should stop the clonazepam that the patient was already on. I *assumed* that a combination of antidepressant medication and CBT would reduce panic, depression, and avoidance and that the benzodiazepines would be stopped. However, the patient said that the clonazepam had been helpful. The patient might feel worse if it was stopped, and stopping it might be associated with some rebound or withdrawal effects. I decided to continue it for another week. I told the patient that I would consider stopping it or cutting down at the next visit.

Dizziness

Dizziness is a common symptom of patients with panic, which may represent some underlying disturbance in the vestibular system (Tecer et al. 2004) and can be related to hyperventilation or occur from other, unknown reasons. Dizziness can also represent a number of medical problems. The patient's medical evaluation had focused on his cardiovascular but not ear, nose, and throat (ENT) system. His interpretation of the symptom was that he might have a brain tumor, and if not that, some other serious medical problem. Should I recommend further medical evaluation? The patient felt that his worst symptom was the one that seemed to interfere the most with his work and that it served as a major cue for his feeling like something was seriously wrong. I weighed the advantages and need of further medical workup against the effect that such a workup could have on reinforcing his sense that he had a "medical problem." An important aspect of care, and implied in the guideline, is the need to help patients understand how their symptoms might have a psychological rather than medical origin. I also assumed that treatment of the panic would result in reduction in dizziness. The dizziness would be used as part of in vitro exposure therapy and in helping the patient learn new breathing techniques. I decided not to make a referral at this time but to follow up if the dizziness did not diminish as expected.

Onset/Stressors

As is often the case, the reason for the onset of the panic was not clear. Stress at work appeared to be a contributing factor, although the patient had had many times at work when he felt under intense stress without feeling panicky. There were no other obvious lifestyle habits (such as caffeine use, cessation of smoking, drug use) or medical precipitants. I wondered if stress management might be useful at some point, but the patient was not interested. The guideline does not recommend stress management for treating panic, but there appeared to be another reason to offer it, that is, for reducing stress and improving coping.

Avoidance

The patient had developed significant avoidance that could eventually cost him his job. I was confident that exposure therapy would reduce his avoidance. I considered this to be my number one goal (the patient was more concerned about his dizziness.) As suggested in the guideline, "massed" exposure, that is, intensive exposure over a short time, can be used and appears as effective as more prolonged exposure. (The focus would be on driving and crossing bridges.) My schedule does not permit providing massed exposure and it is not available in the community, so I decided against using this potentially rapid treatment of his avoidance. I also felt it was important to first develop a relationship with

the patient, to help reduce the dizziness, at which point he might be open to more demanding interventions such as exposure therapy.

Interpersonal Isolation

The patient reported that he had a lot of friends but no close relationships. He said that he had thought about getting married and at one point was engaged, but his then-girlfriend had moved from the area and he didn't pursue the relationship. He spent weekends running in the mountains with a group of other runners. He would sometimes go to a bar with his coworkers but otherwise spent most of his time at home. I felt that his social isolation was an important aspect of his life that should be addressed; however, he was not interested in addressing it, at least initially.

With these considerations, I discussed an initial treatment plan and approach with the patient. I mentioned that I would periodically monitor how he was doing, but I didn't give him the same time frame of improvement reflected in Table 17–1; I only mentioned that he should be feeling better in a few weeks. He agreed to my initial treatment plan as follows: I prescribed 1) sertraline one-quarter tablet for 3 days, then half a tablet, explaining the side effects and telling him to call me if his mood got worse and to stop the medications; and 2) use of a relaxation tape three times a week. I scheduled a visit in 1 week.

Course

Visit 2 (Week 2)

The patient said that he was feeling less anxious and had not had a full-blown panic attack although he had a few episodes of feeling panicky. His mood had improved slightly. He kept records of how he was feeling and brought them in for review. He practiced relaxation as prescribed. He was still avoiding going to work. I asked him to decrease his clonazepam, I increased his sertraline to 25 mg, and I asked him to continue use of his relaxation tape. Much of the session was spent in discussing the nature of panic, as per CBT (Barlow and Craske 2006). I printed out and gave him the National Institute of Mental Health information handout on panic disorder. I considered using a structured treatment program (e.g., Barlow and Craske 2006, which we keep on hand and sell to patients at cost), but I didn't feel it was necessary because the patient was already self-monitoring and I could provide the treatment. I also did in vivo exposure for dizziness and taught him breathing exercises (per Craske and Barlow 2006).

Visit 3 (Week 3)

The patient reported depression, dizziness, and panic similar to the previous week. I continued to provide CBT/psychoeducation and increased sertraline to 50 mg. I continued in vivo exposure and prescribed relaxation training and breathing practice. The in vivo exposure followed Craske and Barlow (2006) and included spinning in his chair while monitoring symptoms, feelings, and cognitions. I also began graduated in vitro exposure focusing on driving and crossing bridges.

Visit 4 (Week 5)

The patient felt depressed on increased sertraline. The patient called the on-duty resident, who recommended that he decrease sertraline to 25 mg and restart clonazepam, which he did. The patient was feeling better and maintained sertraline at 25 mg and clonazepam 0.25 mg prn. I decided to begin bupropion (25 mg) because I felt the patient continued to have depressive symptoms and I was reluctant to go up on the sertraline. If I followed the NICE guidelines, I should have considered imipramine or clomipramine as the next choice and even stopped sertraline. Why did I deviate? I was worried about anticholinergic side effects from tricyclics in a patient with very high somatic concerns. I was also concerned about risk of an overdose and I wanted to use a drug with different pharmacological action. However, there is only anecdotal evidence of the use of bupropion and panic, and it would be an off-label use (Simon et al. 2003).

Visit 5 (Week 7)

The patient said that he felt close to normal. No panic or anxiety, and he was now driving to work on his own. We set a goal of his flying back East to visit his family.

Visit 6 (Week 10)

The patient said he continued to feel some anxiety in the morning (mainly muscle tension). He felt a little more panicky when driving over the bridge but had no panic attacks, and he was not feeling depressed. We reviewed his "exposure work" (assignments to

TABLE 17–1. Initial treatment goals, measures, and methods

Treatment goal	Measure	Method
Reduce panic attack frequency to 0 within 2 months	Self-report of weekly panic attacks	Pharmacotherapy Psychoeducation/cognitive-behavioral therapy
Reduce dizziness from 10 to 3 within 3 months	Self-report: 10=as bad as it gets, 0=no dizziness	Pharmacotherapy Cognitive-behavioral therapy with emphasis on in vitro exposure, breathing training, and relaxation
Reduce depression by 50% or more, by 8 weeks	QIDS-SR_{16}	Pharmacotherapy
Reduce avoidance, return to driving to work within 2 months	Self-report	Exposure therapy would be a key part of the cognitive-behavioral therapy
Reduce stress at work	Self-report	Deferred, per patient preference

Note. QIDS-SR_{16}=Quick Inventory of Depressive Symptomatology–Self-Report.

drive to work, cross bridges, etc.) in terms of feelings/symptoms, behaviors, and cognitions.

Visit 7 (Week 15)

The patient continued to do well with no depression, minimal anxiety, and no avoidance. He continued to drive to work. He had bought a ticket to go back East but was worried about having a panic attack on the plane. I gave him instructions on breathing and suggested he could use clonazepam if he was experiencing significant anxiety before the flight. I asked if he wanted to address the issue of dating, and he agreed but seemed ambivalent and did not want to set any goals related to dating at this time.

Visit 8 (Week 18)

The patient continued to do well with minimal anxiety, depression, and no panic. He decided he did not want to work on dating or on stress issues at work. We discussed moving into maintenance/relapse prevention after the next session.

Visit 9 (Week 22)

The patient continued to do well. We discussed relapse prevention techniques and scheduled a follow-up visit to review medication use.

Visit 10 (Week 38)

This was the maintenance visit. The patient continued to do well with no panic, depression, or anxiety. He wanted to stop medications. I explained that the guideline recommends that he continue medication for 6 months after he has been symptom free and explained the rationale. He was willing to continue for another 3 months. He was informed of how to stop the medications and alerted to discontinuation effects.

Relapse

About 9 months after the last visit, the patient called to say that he had had another panic attack out of the blue. He said that about 3 months after his last visit he stopped the bupropion, and because he continued to feel very good, he also stopped the sertraline. About 1 week prior to his call, he had another panic attack and decided to restart the sertraline and clonazepam on his own. One day after restarting the medication he became depressed and frightened. He was worried that he would fall right back into his old patterns. I scheduled a visit within the next week, suggested he continue the sertraline but at a lower dose, and invited him to call me if he felt worse. I reminded him of the importance of his breathing exercises, which he had stopped. At our session a week later, he reported having had no panic attacks and his mood was much improved (3/10; maximum dysphoria 4/10). Surprisingly, he had no dizziness. He was not doing his breathing or relaxation but was not avoiding. I suggested he continue the sertraline but decrease the clonazepam. In this instance, the onset seemed to be related to a significant increase in his workload and the feeling that he was falling behind and not doing a good job. However, he did not want to discuss these issues in detail. Based on his relapse, longer-term use of the medication (at least another year) seemed advisable. For the panic,

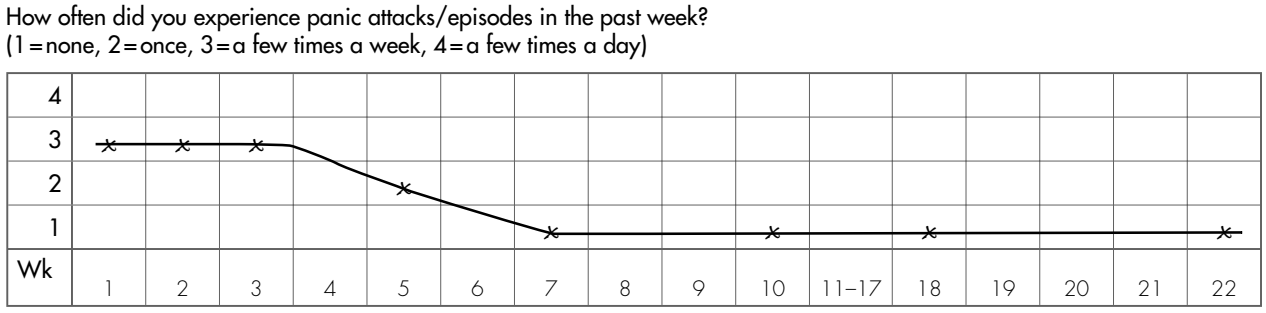


FIGURE 17–1. Panic attacks/episodes.

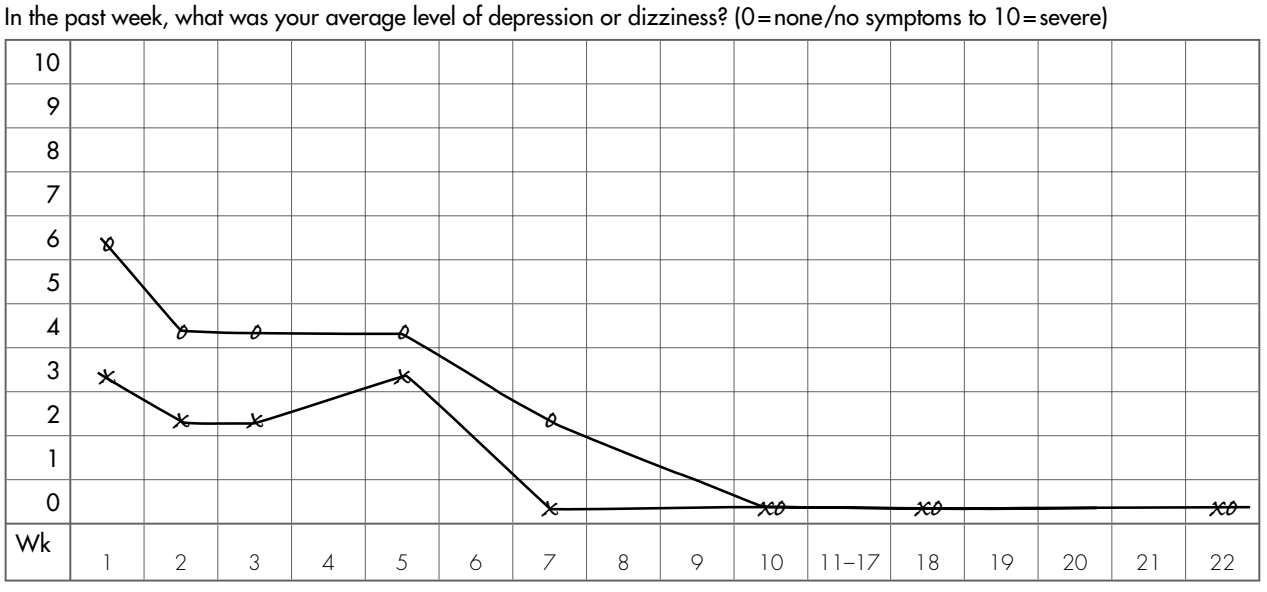


FIGURE 17–2. Depression (X–X) and dizziness (O–O) trends.

stress reduction/coping skills might have reduced his stress at work.

Figures 17–1 and 17–2 show the patient's progress over the first 22 weeks of therapy.

Summary of Guideline Use

Table 17–2 summarizes the guidelines applied to this patient and whether or not the guideline was followed. Although an SSRI was used as the first medication of choice, the actual practice was more complicated because the patient had had an adverse reaction to the first drug prescribed in that class and a second SSRI was started at a unusually small dose. Benzodiazepines were prescribed by the patient's internist and were continued for several weeks, which would seem to go against the spirit of the guideline. The guideline stresses that "CBT should be delivered only by suitably trained and supervised people who can demonstrate that they adhere closely to empirically grounded treatment protocols." Treatment protocols are available for panic disorder (and many other problems), for instance, through the Treatments That Work series (www.oup.com/us/catalog/general/series/TreatmentsThatWork/?view=usa).[1] Since I have extensive experience in treating panic, it didn't seem necessary to strictly follow such a protocol. The guideline also expects CBT to be com-

[1] The author is an unpaid member of the Treatments That Work Scientific Advisory Board.

TABLE 17–2. NICE guideline recommendations for panic disorder

Area	Recommendation	Followed
Shared decision making	Include review and benefit of therapies, alternative therapies	√
Patient education	Provide information on nature, course, and treatment, including medication use and side effects	√
	Use everyday, jargon-free language	√
Medication	*Step 1:* SSRIs are medications of first choice	√
	Step 2: If an SSRI is not suitable or there is no improvement after a 12-week course and if a further medication is appropriate, imipramine or clomipramine	No, we preferred bupropion
	Benzodiazepines are not to be used	No, patient was already on benzodiazepines
	Sedating antihistamines or antipsychotics should not be used	NA
	Consider medication in terms of age, previous treatment response, risk, likelihood of overdose/self-harm, tolerability, patient preference, cost (where equal effectiveness is demonstrated)	√
	Doses at the upper end of the indicated dose range may be necessary and should be offered if needed	NA
	Long-term treatment may be necessary for some people and should be offered if needed	NA
	If the patient is showing improvement on treatment with an antidepressant, the medication should be continued for at least 6 months after optimal dose, then the dose can be tapered	√
	Patient should be advised to take medication as prescribed, discontinuation/withdrawal symptoms should be discussed	√
Psychotherapy	CBT is therapy of first choice	√
	Adherence to CBT treatment guidelines	No, adapted to patient needs
	7–14 hours over 12–16 weeks or brief CBT (6–8 hours) with self-help	9 hours over 22 weeks without self-help
	Intensive "massed exposure"	Not available
	Relapse prevention	√
Self-help	Bibliotherapy based on CBT should be offered	√, NIMH sheet given
Support groups	Information on support groups should be offered	Not needed
Exercise	The benefits of exercise should be discussed	Already exercising
Monitoring	Short self-report questionnaires can be used	I used Likert-like scales
	Use psychometrically sound measures as appropriate	
	Baseline assessment	√
	Process assessments	√
	12-week review	Done at 10 weeks

Note. CBT=cognitive-behavioral therapy; NA=not applicable to patient; NICE=National Institute for Health and Clinical Excellence; NIMH=National Institute of Mental Health; SSRI=selective serotonin reuptake inhibitor.

Source. Adapted from www.nice.org.uk/Guidance/CG22 for panic disorder.

pleted in 4 months, although this was not possible with the patient's schedule.

I find that keeping track of the patient's progress is the most useful aspect of the guideline. Medication algorithms are designed so that failure to improve by a certain period leads to other medication choices. This is not the case with psychotherapy, where failure to improve leads to another set of recommendations. I would expect a patient with uncomplicated panic to have significantly improved by six sessions (Kenardy et al. 2003; Taylor 2006). If the patient had not improved, I would ask myself, as noted in the introduction: 1) Is this the right therapy? 2) Is the therapy being followed (e.g., am I doing the CBT properly, is the patient taking the medication?) or 3) Am I missing something? (e.g., wrong diagnosis, a medical problem, secondary gain, the patient's dissatisfaction with me or the therapy). In this case, the patient had improved as expected.

Ways to Improve Practice

In reviewing the guideline and my practice, I identified some areas where I could improve and some questions where I would like more information. I could improve by ensuring that self-help manuals are available when needed (e.g., order a supply in advance, sign up for a supplementary program). Although I am confident in my ability to provide CBT, a manual helps to ensure that the patient is provided the core information. I might have focused more attention on how the patient might cope or avoid relapse.

References

Barlow DH, Craske HG: Mastery of Your Anxiety and Panic: Workbook (Treatments That Work). New York, Oxford University Press, 2006

Craske HG, Barlow DH: Mastery of Your Anxiety and Panic: Therapist Guide (Treatments That Work). New York, Oxford University Press, 2006

Kenardy JA, Dow MGT, Johnston DW, et al: A comparison of delivery methods of cognitive-behavioral therapy for panic disorder: an international multicentre trial. J Consult Clin Psychol 71:1068–1075, 2003

Simon NM, Emmanuel N, Ballenger J, et al: Bupropion sustained release for panic disorder. Psychopharmacol Bull 37(4):66–72, 2003

Taylor CB: Panic disorder. BMJ 332:951–955, 2006

Tecer A, Tükel R, Erdamar B, et al: Audiovestibular functioning in patients with panic disorder. J Psychosom Res 57:177–182, 2004

This page intentionally left blank

18

Major Depressive Disorder and Bulimia Nervosa

C. Barr Taylor, M.D.

Introduction and Setting

See Chapter 17.

Illustration

This study illustrates:

- Use of American Psychiatric Association (APA) major depressive disorder guidelines
- Use of Sequenced Treatment Alternatives to Relieve Depression (STAR*D) guidelines
- Reconciliation of National Institute for Health and Clinical Excellence (NICE) guidelines for bulimia nervosa with APA guidelines for major depressive disorder
- Use of self-help materials
- Stepped care for bulimia; marginal outcome

Chief Complaint

Ms. S.W. is a 24-year-old female Hispanic law student who presents with a chief complaint of depression and frequent vomiting.

Present Illness

Ms. S.W. said that she had been very depressed for the last 3 months following a breakup with her boyfriend. She said that her relationship with her boyfriend had been on and off again for a number of years but after their last breakup he began dating another woman. For the past 2 months she had felt sad nearly every day and had a loss of interest in her usual activities, fatigue, feelings of worthlessness, and trouble concentrating but no change in weight or sleep. She denied thoughts of suicide. Her symptoms had begun to interfere with her being able to concentrate on her class work. She denied any previous episodes of depression. She had seen a therapist about 3 years ago, when she was in college, around issues of her breaking up with another boyfriend.

She reported bingeing two or three times a day with at least one episode followed by vomiting. She said this habit had been going on for about 4 years and that it was her way of controlling her weight. (Her weight, body mass index = 21, was well within the normal range and had not changed since she was an adolescent.) Typically she would overeat or binge at lunch or dinner and then vomit an hour or so after the meal. She would occasionally binge and purge at night as well. When she wasn't able to vomit, for instance, because she was at an after-dinner party in a friend's apartment, she became very anxious, almost frantic.

Other Significant Findings From Assessment

The patient denied tobacco or caffeine use. She exercised every day. She said she sometimes would drink two to three glasses of wine on a weekend. She reported no drug use or other pharmaceuticals (such as over-the-counter diet pills). She saw her general

practitioner 1 month prior to her initial visit. Physical exam and electrolyte levels were normal. She did not have dental erosion or enlarged parotid glands. Her STAR*D Quick Inventory of Depressive Symptomatology—Self-Report ($QIDS\text{-}SR_{16}$) score was 13 (moderate depression). On her baseline assessment, she rated her average level of depression over the past week as being about a 6 or 7 (on a scale of 0=no depression, 10=moderate to severe depression) and her maximum depression about the same. She rated her average anxiety as being a 6 or 7 and maximum at 8. (She did not consider anxiety to be a major problem.)

DSM-IV-TR Diagnosis

Axis I	Major depressive episode
	Bulimia nervosa
Axis II	None
Axis III	None
Axis IV	Breakup with boyfriend
Axis V	GAF score: 75

Treatment Plan Considerations

Depression and Bulimia

The patient's main problems were depression and bulimia. She was much more concerned about her depression than her bulimia but felt that she wanted to address both issues. She felt the depression was entirely due to her breakup with her boyfriend.

Selecting a Guideline

I considered four guidelines and algorithms: 1) the APA guideline for major depressive disorder (www.psychiatryonline.com/pracGuide/pracGuidehome.aspx), 2) the STAR*D guideline (Rush et al. 2003), 3) the APA guideline for eating disorders (www.psychiatryonline.com/pracGuide/pracGuidehome.aspx), and 4) the NICE guideline for treating bulimia (www.nice.org.uk/CG009niceguideline).

The APA general depression guidelines provide an excellent overview and general principles of treatment (see Table 18–1). I added a row to the guideline to indicate the need to assess and address exercise. Clinicians should consider evaluating the evidence for practices they consider useful and adding them to their personalized guidelines. For example, I used the five-step approach discussed in Chapter 2 of this volume to address the question, "Does exercise help treat depression?" The results are summarized in Table 18–2. I felt there was enough evidence to include exercise in this patient's regimen and to add it to my guideline.

An important consideration for Ms. S.W. was whether to use psychotherapy and/or medications. The patient preferred psychotherapy (including cognitive-behavioral therapy [CBT]) but was willing to consider medications. I was confident the patient would improve on either or both. However, I felt that a CBT approach would be necessary for her bulimia nervosa where medication was less likely to be useful. By using both medication and CBT, I felt that I would be able to provide treatment for her depression (with pharmacotherapy) and to spend the psychotherapy focusing on the eating disorder and relationship issues.

The medication algorithm embedded in the APA guidelines for depression is excellent, but given the extensive use of the STAR*D algorithms and the definitions of response and phases, I think the latter is preferable, in part because of its real-world approach (see Chapter 21 for a detailed presentation of STAR*D) (Gaynes et al. 2008). A critical issue in all the guidelines is to define what represents response and what represents remission. STAR*D defines *response* as a significant improvement in depressive symptoms, although residual symptoms may be present. It is generally defined as a 50% improvement or greater reduction from baseline score on a standard measure such as the Hamilton Rating Scale for Depression, Beck Depression Inventory, or $QIDS\text{-}SR_{16}$. Remission is considered full restoration of a patient's normal capacity for psychosocial and occupational function with no residual symptoms, as defined by a score of ≤7 on the 17-item Hamilton Rating Scale for Depression or <5 on the $QIDS\text{-}SR_{16}$. *Relapse* is defined as recurrence of symptoms and impairment following response. *Treatment-resistant depression* is defined as failure to respond to at least two or three antidepressants given at therapeutic doses for more than 4 weeks (see Chapter 21 and Trivedi et al. 2007 for more details on the STAR*D algorithm).

I debated between the APA eating disorder guideline and the NICE guideline. The NICE guideline is more specific than the APA guideline and is closer to my therapeutic approach. Furthermore, members of our clinic have written a patient treatment guide with an accompanying manual (Agras and Apple

TABLE 18–1. Modified American Psychiatric Association general recommendation for major depressive disorder through remission

Recommendation		Followed
Psychiatric management	Basic clinical requirements	
	Diagnostic evaluation (including safety, functional impairment, etc.)	√
	Establish and maintain a therapeutic alliance	√
	Monitor transference/countertransference	√
	Monitor status and safety	√
	Provide education to patient	√
	Provide education to families	No, per patient
	Enhance medication treatment adherence by emphasizing:	
	When and how often to take	√
	Typical 2- to 4- week lag to benefit	√
	Continue medication even when feeling better	√
	Consult with medical doctor if problems arise	√
	Review side effects	√
	What to do if problems arise	√
	Address early signs of relapse	√
Acute treatment	Select initial treatment modality	
	Antidepressant medication or	
	Psychotherapy or	
	Psychotherapy plus antidepressant medication and	√
	Consider ECT	NA
	CBT and interpersonal therapy have best-documented efficacy	
	Must be integrated with psychiatric management and other treatments	NA
	Response to therapy should be carefully monitored	√
	Care needs to be coordinated	NA
	Patients should be treated until a complete response occurs. If moderate improvement or less is not achieved in 4–8 weeks, treatment needs to be reviewed.	√, used STAR*D criteria
	No or partial response (change medications; add or change psychotherapy; consider ECT)	√
	Reassess in 4–8 weeks	√
	The benefits of exercise should be discussed and program initiated if appropriate[a]	*Already exercising*
Continuation phase (16–20 weeks following remission)	Patients should continue on medications, generally at the dose used in the acute phase	√
	Psychotherapy effective for relapse prevention should be used (frequency of visits is not clear)	NA

Note. CBT=cognitive-behavioral therapy; ECT=electroconvulsive therapy; NA=not applicable to patient; STAR*D=Sequenced Treatment Alternatives to Relieve Depression.

[a]Author's addition.

Source. http://www.psychiatryonline.com/pracGuide/pracGuidehome.aspx.

TABLE 18–2. Using the 5-step evidence-based medicine process to evaluate the effects of exercise on depression

Process	Evaluation
Step 1. Formulate the question	Does exercise help treat depression? (If yes, other important issues would need to be considered, such as type, intensity, frequency, and setting.)
Step 2. Search for answers	Several reviews suggested that regular aerobic exercise is effective in reducing depressive symptoms (e.g. Barbour et al. 2007) but that it is not recommended as a primary treatment, at least in older adults (Steinman et al. 2007). An older meta-analysis (Lawlor and Hopker 2001) concluded, "The effectiveness of exercise in reducing symptoms of depression cannot be determined because of a lack of good quality research on clinical populations with adequate follow up."
Step 3. Appraise the evidence	The advantage of searching for reviews and meta-analyses is that others have appraised the evidence. The results and conclusions are not always consistent, however, as suggested in the two reviews above, and I am left to draw my own conclusions. I don't find convincing evidence that exercise should be the primary treatment for depression. Having been involved in a number of exercise intervention studies, I am also aware of the time and resources required, the potential medical risks (falls, sprains, cardiovascular events), and the generally poor adherence. I conclude, as reflected in Table 18–1 (which comes not only from this review, but my own work in the field) that the benefits of exercise should be discussed and a program initiated if appropriate.
Step 4. Apply the results	Ms. S.W. was already exercising. (My general approach to provide exercise programs for patients is presented in Miller and Taylor 1995).
Step 5. Assess the outcome	The APA guideline includes periodic assessment of depression, and I use frequent single-item assessments. However, it is often difficult to attribute change to any one practice when patients are being provided a number of activities. I think it is important to assess adherence to exercise, and I routinely ask patients to report how many times a week they exercise as a global assessment.

2008a, 2008b) that can be used to provide the CBT approach recommended by NICE.

The NICE guideline for the pharmacological treatment of bulimia nervosa is not consistent with the STAR*D and APA guidelines for major depressive disorder. The NICE guideline (Table 18–3) states that fluoxetine is the medication of first choice; STAR*D recommends citalopram (although probably any selective serotonin reuptake inhibitor could be substituted). NICE states that the patient should be told to expect rapid improvement with pharmacology (based on a limited number of studies); APA emphasizes the need to inform patients of more gradual improvement. NICE recommends reassessment at 4–5 months (based on expected time course for bulimia treatment); the STAR*D guidelines suggest reassessment at 8 weeks.

Another consideration was what assessment instruments to use for bulimia nervosa. Although a number of standardized instruments are available to assess eating disorder attitudes and behaviors, I decided to focus only on purging episodes as the main outcome measure.

Relationship Issues

The patient's depression was related to the breakup of a relationship, and this would be a focus of therapy. I am not aware of guidelines for treating relationships, although therapeutic practices of relevant interventions such as interpersonal therapy (IPT) have been clearly defined. I planned to use IPT techniques to deal with the relationship issue and CBT techniques to deal with the bulimia and depression. Many of the core ideas of these therapies cross pa-

TABLE 18–3. Modified NICE guidelines for bulimia nervosa

Recommendation		Followed
Psychological interventions	As a first step, encourage use of evidence-based self-help program	No, given manual
	If self-help is declined, offer CBT	
	16–20 sessions over 4–5 months	√
	Reassess at 4–5 months	√
	If CBT is declined or no response (per guideline), offer other psychological treatments. IPT should be considered as an alternative to CBT but patient should be told it may take 8–12 months.	±
Pharmacological interventions	As an alternative or additional first step, offer antidepressant	√
	Fluoxetine is first choice. Dose for bulimia nervosa is higher than for depression.	Used STAR*D
	Inform patients that long-term effects are not known; beneficial effects will be rapidly apparent	
Medical management	If patient is vomiting frequently or taking many laxatives, assess fluid and electrolyte balance	No
	If electrolyte disturbance is detected, focus on behavior change or oral supplementation	NA

Note. CBT=cognitive-behavioral therapy; IPT=interpersonal therapy; NICE=National Institute for Health and Clinical Excellence; NA=not applicable to patient; STAR*D=Sequenced Treatment Alternatives to Relieve Depression.

Source. http://www.nice.org.uk/CG009niceguideline.

tient issues. For instance, negative thoughts related to body image are likely to cause both depression and excessive focus on weight and shape concerns. The NICE guideline for eating disorders states that if IPT is used, the patient should be cautioned that it might take up to a year. But with any recommendations derived from trials that restrict therapy to one approach, the recommendation does not fit the real world where we use several synergistic approaches.

Another issue was how to measure progress in the patient's relationships. Relating involves a number of factors, including confidence, skills, interest, and behavior. I decided to use a very simple metric by asking the patient every few sessions how satisfied she was with her relationships, from 1=not at all to 4=very satisfied, with the idea that I could explore the components as it seemed appropriate. I assumed she would need to overcome her sense of loss and impaired self-esteem from the breakup with her boyfriend before she would want to explore and develop new relationships. Routine use of even this simple question would remind me to keep relationship issues as part of the agenda, because this was the patient's main concern, and also allow me to determine if she was making progress. From my clinical experience I would expect her relationships to improve to a score of 3 or 4 after 6 months of therapy.

With these considerations, I discussed an initial treatment plan and approach with the patient. I mentioned that I would periodically monitor how she was doing. She agreed to my initial treatment plan, listed in Table 18–4.

I prescribed citalopram 10 mg, increasing to 20 mg as tolerated. I explained the side effects and told her to stop the medication if she got worse. I scheduled a visit in 1 week.

Course

Visit 2 (Week 2)

The patient had no side effects from the citalopram and had increased it to 20 mg per day. About half the session was spent reviewing her relationship with her ex-boyfriend. It appeared that he had pushed

TABLE 18–4. Initial treatment goals, measures, and methods

Treatment goal	Measure	Method
Reduce depressive symptoms by 50% by 8 weeks	QIDS-SR_{16}	Pharmacotherapy Supportive therapy
Reduce bulimic episodes by 50% by 8 weeks	Self-report (episodes/week)	Cognitive-behavioral therapy Pharmacotherapy?
Deal with relationship issues	Self-report (qualitative)	IPT embedded in cognitive-behavioral therapy Supportive psychotherapy

Note. IPT = interpersonal therapy; QIDS-SR_{16}=Quick Inventory of Depressive Symptomatology–Self-Report.

her away as she was trying to achieve a closer relationship. She interpreted this as "I am not worthy or attractive" and "Nobody will ever want me"—thoughts appropriate for CBT intervention. The second part of the session was spent providing an overview of CBT treatment for bulimia nervosa. She was provided a self-help patient treatment manual, *Overcoming Eating Disorders* (Agras and Apple 2008a). According to the therapist guide (Agras and Apple 2008b), following evaluation, for the first session I should have undertaken the steps listed in Table 18–5. I followed most of these activities although I knew I would need to have some flexibility in scheduling.

Visit 3 (Week 3)

The patient was feeling less depressed and had no side effects from the medication. She had not completed her daily food records or read the first chapter of the workbook. We spent about 30 minutes on relationship issues, which was what she wanted to talk about. We then returned to discuss the importance of keeping the food records and reading the first several chapters of the manual. In reviewing the therapist guide (Agras and Apple 2008b), I found there was little said about very noncompliant patients, although much has been written about this in the CBT literature. I reemphasized the importance of keeping the records and examined her thoughts about doing so. The patient said that she didn't have time and felt too conspicuous carrying them around. We problem-solved ways for her to keep the records (such as writing the items down before she went to sleep) and even focused on accurate record keeping for a few days but not the entire week. I had the impression the patient was not motivated to change her bingeing/purging. I decided to wait another week or so to confront, giving me more time to build a therapeutic relationship.

Visits 4–8 (Weeks 4–12)

The patient's mood had improved. She no longer met criteria for depression by week 8 and would be considered in remission. She was feeling much better about herself in general except for her bingeing/purging. It was evident to her that her bingeing/purging was a factor in her break-up with her boyfriend and would be a factor in future relationships, but she was not motivated to stop. She firmly believed that her bingeing/purging was controlling her weight. Overall adherence to the CBT practices was less than 50%. She was not "able" to follow a prescribed diet plan. Although her binge frequency, by her report, had dropped by 25%–50% and she had days when she was binge/purge free, I felt it was important for me to review why the patient was not doing better. I didn't feel her diagnosis or assessment was wrong, but she had very little motivation to change her behavior or to question her core beliefs and attitudes about bingeing/purging. We reviewed treatment options, including whether she might want to see someone else or change or increase her antidepressant, all of which she declined.

Weeks 13–36

The patient remained in remission from her depression for the next 4 months and wanted to stop her medication. She was feeling much better about herself and her relationships and had started dating. However, she continued to binge/purge at a high rate. I did a literature search to see if there were some new approaches that might prove beneficial, following the five-step approach discussed in Chapter 2 of this volume (see Table 18–6). The search

TABLE 18–5. First session of cognitive-behavioral therapy for bulimia nervosa, therapist activities

Activity	Followed	Comment
Provide rationale for therapy	√	
Emphasize importance of regularizing eating patterns	√	
Explain the three phases of treatment (behavior change, binge triggers, relapse prevention)	√	
Outline the session structure	No	I planned to focus on several issues, including medication and relationships
Provide evidence base of outcome	√	
Provide rationale for self-monitoring and introduce the daily food records	√	
Introduce the patient workbook	√	
Assign homework	√	Read first chapter in workbook, and complete daily food records

Source. Agras and Apple 2008b.

identified topiramate as a possible adjunct, but I felt the evidence for the use of topiramate was marginal. I discussed the risk/benefit with the patient, and she wanted to try it. After several weeks, and at a dose of 150 mg, the patient had seen no benefit and asked to stop the medication.

We stopped therapy after 36 weeks. The patient had accepted a job as an intern at a law firm in New York for the summer, and we decided this was a good time for her to take a break from therapy with the notion we would reevaluate treatment options in the fall when she returned, if she wanted to.

TABLE 18–6. Using the 5-step evidence-based medicine (EBM) process to evaluate pharmacological approaches to reducing bingeing/purging

Process	Evaluation
Step 1. Formulate the question.	Are there medications that can be used to reduce bingeing/purging (aside from those discussed in guidelines)?
Step 2. Search for answers.	After a more general search, I found some reports of the potential benefits of topiramate (Kotwal et al. 2003; Nickel et al. 2005). I could not find the Kotwal article online through my medical library (limited access to many journal articles is a major problem for EBM in clinical practice). Nickel et al. 2005 found that about 30% of the topiramate group had reduced binge frequency by 50% or more, compared with 3% in the control group. The medication was started at 25 mg and increased to 250 mg over 6 weeks. No side effects were noted. I also reviewed the basic pharmacology of topiramate.
Step 3. Appraise the evidence.	The evidence for the use of topiramate was marginal. I discussed the risk/benefit with the patient, however, and she wanted to try it.
Step 4. Apply the results.	Ms. S.W. was started on topiramate at 25 mg, with the goal of increasing the dose to 250 mg over 6 weeks.
Step 5. Assess the outcome.	After 4 weeks, and at a dose of 150 mg, the patient had seen no benefit and asked to stop the medication.

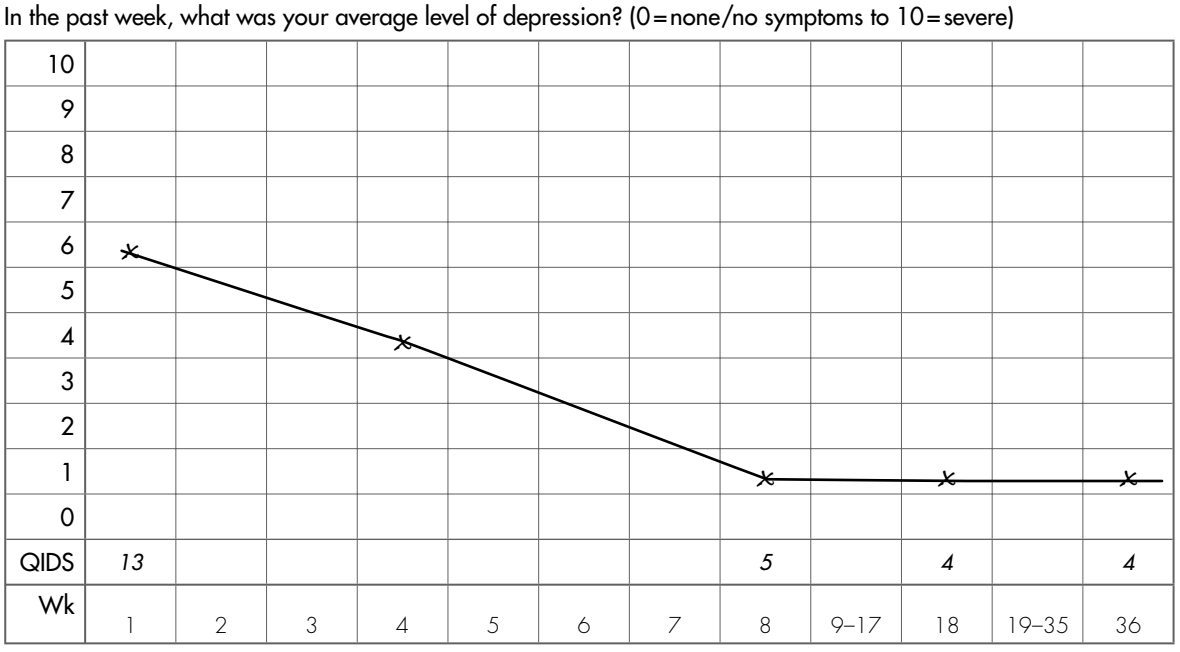


FIGURE 18–1. Depression trends (X–X).
QIDS=Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}).

Figures 18–1 through 18–3 show the patient's progress in mood and depression, purging, and relationships over the first 36 weeks of therapy. Her depression and her relationships improved; there was little change in her purging.

Summary of Guideline Use

At the end of 36 weeks I reviewed my use of the guidelines. I followed the APA guidelines for depression fairly well (see Table 18–1). The guidelines suggest that the therapist review transference/countertransference issues. While I hope I do this routinely, I think I became more invested in the patient not being bulimic than she was. I was frustrated that she had not responded as anticipated to the manualized CBT approach. The patient responded to the first step of the STAR*D guideline. However she did not want to continue her medication for a full 12 months.

The NICE guideline recommends self-help as the first step and I provided the patient with a manual, but since she was in therapy it seemed appropriate to provide additional therapeutic guidance. I was not very successful at implementing a course of CBT for bulimia via a manualized treatment. However, I was able to have the patient use the basic skills

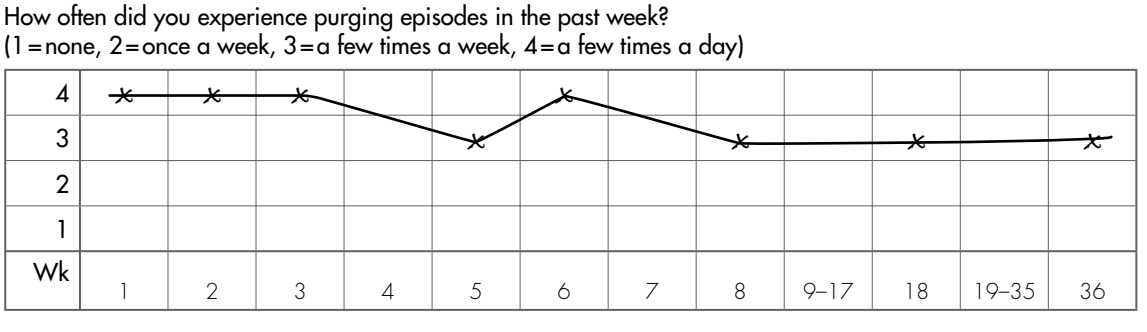


FIGURE 18–2. Purging episodes.

FIGURE 18–3. Relationship satisfaction.

(food monitoring, meal planning, identifying binge triggers) by tailoring the approach to her preferences. The NICE guideline suggests that I might have moved from CBT to IPT, if CBT was unsuccessful. However, a focus of IPT would have been on her relationships, which had improved substantially. The NICE guideline also suggests that IPT needs to be provided for a year, but I didn't think this was justified. The NICE guideline says that no medications other than fluoxetine should be used. As mentioned, I decided to use citalopram following STAR*D and was not convinced that fluoxetine was required. A PubMed search identified topiramate as a possible adjunct. Alternatively, I could have increased the dose of citalopram. The NICE guideline recommends review of electrolytes, which I failed to do. The patient's physical exam and electrolyte levels were normal at baseline; a reassessment would have been appropriate and would also have provided an indication of my concern about the potential medical consequences of her behavior. I believed I had generally followed the guidelines and a stepped-care approach with benefit for her depression, self-image, and relationship issues but with little impact on her bingeing.

Ways to Improve Practice

I identified several aspects of my practice that might have benefited this patient. First, I might have been more conscientious in how I delivered the manualized treatment. I might have put greater emphasis on the need for her to use it and examined her resistance more effectively. Second, I might have presented this case to the clinical case conference where other experts could present ideas relevant to the patient's care. This would also have given me an opportunity to examine the therapeutic relationship. Finally, I should have monitored her electrolytes periodically to help emphasize the potential negative effects of her persistent purging.

References

Agras WS, Apple R: Overcoming Eating Disorders: A Cognitive-Behavioral Treatment for Bulimia Nervosa and Binge-Eating: Patient Manual (Treatments That Work). Oxford, UK, Oxford University Press, 2008a

Agras WS, Apple R: Overcoming Eating Disorders: A Cognitive-Behavioral Treatment for Bulimia Nervosa and Binge-Eating—Therapist Guide (Treatments That Work). Oxford, UK, Oxford University Press, 2008b

Barbour KA, Edenfield TM, Blumenthal JA: Exercise as a treatment for depression and other psychiatric disorders: a review. J Cardiopulm Rehabil Prev 27:359–367, 2007

Gaynes BN, Rush AJ, Trivedi MH, et al: The STAR*D study: treating depression in the real world. Cleve Clin J Med 75:57–66, 2008

Kotwal R, McElroy SL, Malhotra S: What treatment data support topiramate in bulimia nervosa and binge eating disorder? What is the drug's safety profile? How is it used in these conditions? Eat Disord 11:71–75, 2003

Lawlor DA, Hopker SW: The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. BMJ 322:763–767, 2001

Miller NH, Taylor CB: Lifestyle Management for Patients With Coronary Heart Disease. Champaign, IL, Human Kinetics, 1995

Nickel C, Tritt K, Muehlbacher M, et al: Topiramate treatment in bulimia nervosa patients: a randomized, double-blind, placebo-controlled trial. Int J Eat Disord 38:295–300, 2005

Rush AJ, Trivedi M, Fava M: Depression, IV: STAR*D treatment trial for depression. Am J Psychiatry 160:237, 2003

Steinman LE, Frederick JT, Prohaska T, et al: Late life depression Special Interest Project (SIP) panelists: recommendations for treating depression in community-based older adults. Am J Prev Med 33:175–181, 2007

Trivedi MH, Rush AJ, Gaynes BN, et al: Maximizing the adequacy of medication treatment in controlled trials and clinical practice: STAR(*)D measurement-based care. Neuropsychopharmacology 32:2479–2489, 2007

19

Depression, Refusal to Eat, and Nasopharyngeal Carcinoma

C. Barr Taylor, M.D.

Introduction and Setting

See Chapter 17.

Illustration

This study illustrates:

- Application of guidelines to psychiatric and medical comorbid problems
- Monitoring of symptoms
- Emergence of new problems
- A threat to the therapeutic relationship

Chief Complaint

Ms. P.A. is a 60-year-old woman who presents with a chief complaint of depression and being unable to eat.

Present Illness

The patient said she was fine until 3 months prior to admission, when she was diagnosed with nasopharyngeal carcinoma. She began radiation therapy followed by chemotherapy that needed to be stopped after two rounds because of side effects. About 6 weeks prior to admission she began to feel very depressed and was started on citalopram 20 mg but felt that the medication only had a mild benefit and caused her mouth to be drier. She said that food has no taste and that eating foods was very painful. She reported feeling sad most of the day, had no appetite or motivation to eat, had trouble concentrating, used sleep to escape, had no interest in sex, and had fatigue and no energy. She denied any suicidal intent. She reported attending a group for cancer survivors, seeing an acupuncturist for pain control, and using several compounds prescribed by a Chinese herbalist. She would sometimes use the dietary supplement Ensure to increase her caloric intake.

Past Psychiatric History

Ms. P.A. said she was a recovering alcoholic and had been abstinent for over 5 years. She attended Alcoholics Anonymous (AA) for 2 years after she quit drinking but had not been to AA since.

Medical History

The patient underwent radiation and chemotherapy for nasopharyngeal carcinoma. No nodes were found, and neck dissection was not recommended. Her weight dropped from 140–150 lb prior to cancer treatment to her current weight of 114 lb. She was told that she would need to have a feeding tube if she did not start to gain weight.

Social History

The patient was married for 30 years and had two adult children (both women). She worked at home as a part-time accountant. The patient felt that her cancer treatment had had a very negative impact on her family and marriage. She had lost interest in sexual activity and worried that her husband was very

frustrated (she had asked him about this and he didn't complain) and that she had let down her daughters by not being stronger in simply eating "at any cost."

Other Significant Findings From Assessment

The patient used to play tennis two to three times a week and play golf but now reported walking three to four times a week for 30 minutes. She denied tobacco or caffeine use. On the mental status examination, she was tearful. There was no evidence of cognitive impairment. She scored 17 (moderate to severe depression) on the Quick Inventory of Depressive Symptomatology–Self-Report ($QIDS\text{-}SR_{16}$); 4/10 on the self-report of depression; and 1/10 on anxiety. Her Mini-Mental State Examination score was 29.

DSM-IV-TR Diagnosis

Axis I	Major depressive episode History of ethanol abuse
Axis II	None
Axis III	Radiation and chemotherapy for nasopharyngeal carcinoma with secondary pain upon swallowing, loss of taste
Axis IV	Stress/trauma related to cancer, possible marital discord
Axis V	GAF score: 60

Treatment Plan Considerations

Depression

The patient presented with major symptoms of depression probably related to the effects of being diagnosed and treated for cancer, with loss of taste and pain on swallowing, resulting in reluctance to eat and weight loss. She felt that she had let her family down because she was not eating. She was desperate to begin eating because she was terrified of having a feeding tube placed. The use of unidentified compounds from the Chinese herbalist presented a problem of potential adverse interactions among their use and psychopharmacology. (I made a mental note to do a quick search on the use of antidepressants and herbs and Chinese medicines when I had time and told the patient to try to find the names of the compounds she was using.)

Selecting a Guideline

As in the case presented in Chapter 18, the American Psychiatric Association (APA) major depressive disorder guideline would be appropriate for general management and the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) guideline would be appropriate for medication (see Chapter 21 for a detailed presentation of STAR*D). The STAR*D guideline would call for 8 weeks of citalopram at an adequate dose; the patient had been at an adequate dose for 6 weeks. She also felt that her mouth was drier on citalopram than it had been before she started the medication, but ensuing rounds of radiation and chemotherapy are likely to have reduced her salivation. STAR*D suggests sertraline, venlafaxine extended release, or bupropion sustained release as an alternative if citalopram is not effective or is unacceptable. I was concerned that sertraline might cause more drowsiness without improving salivation; I decided to use fluoxetine based on my clinical impression that it can be more activating and has little effect on salivation (evidence: soft and clinical). (The use of monoamine oxidase inhibitors, or MAOIs, to be considered if the patient did not respond to the first three steps, would be contraindicated by her use of meperidine before eating to reduce the pain of swallowing.) As noted in the APA guideline, coordination of care with her other treatment providers would be important.

The APA guideline for depression would suggest that cognitive-behavioral therapy (CBT) and interpersonal therapy would be appropriate. I felt that a very structured, behavioral approach would be needed to help her with eating. I also felt that supportive therapy would be most appropriate following the general principles discussed in Winston et al. (2004).

Weight Loss/Food Refusal

The patient's main goal was to resume eating. Her problems in eating were probably related to her depression (loss of appetite) but also to her loss of taste and pain on swallowing. It hurt to eat. There are no guidelines for food refusal that I could find. We decided to establish two measurement parameters: weight gain and difficulty swallowing. I assumed, based on my experience with cancer patients who had gone through this type of treatment, that swallowing would improve, but gradually and subtly. I

felt that the patient would benefit by seeing progress on a graph, even if slow. If progress didn't occur, it would help to consider other treatment approaches. A note in her medical chart indicated that she had been prescribed 45 mg of zinc sulfate tablets three times per day with food "in order to potentially provide some benefit to her taste function as previously described in the literature."

Stressors/Trauma of Cancer

A number of clinical approaches to treating cancer trauma have been described (Spiegel and Classen 2000). I have had positive experiences with online groups as one method of providing this service (Winzelberg et al. 2003). Encouraging the patient to use Internet resources, such as searching for ways to help treat her pain and improve her swallowing, would also seem appropriate.

Relationships Issues

The patient felt that her medical problems and depression had adversely affected her relationships. I decided to provide supportive therapy as noted above to help her with these problems and to consider referral to couples therapy (see Table 19–1).

I prescribed fluoxetine 20 mg one tablet in the morning and the use of a relaxation tape three times a week. I also gave her an individualized food record and asked her to write down everything she ate and, for at least one meal a day, what her thoughts were around eating (e.g., "I want my food to taste better," "I will never get better"). I also encouraged her to explore foods that might be easier to swallow and relatively more satisfying. I scheduled an appointment in a week.

Course

Visit 2 (Week 2)

The patient reported progress in being able to eat and found the logs useful. She was proud of herself when she could eat and hard on herself when she was not able to. She found that her mood varied throughout the week and at times she "felt normal" and other times felt frustrated and down. Evening meals were much harder for her. She was not preparing meals for her family, in part because she resented the fact that she couldn't eat what she prepared. She decided it would be important for her self-esteem if she began to prepare meals even if she wasn't able to eat them. The patient's depression scores, average level of eating/swallowing difficulty, and weight can be seen in Figures 19–1 and 19–2.

Visit 3 (Week 3)

The patient had lost another pound and was feeling discouraged. She realized that she would never be able to gain weight or even maintain her weight unless she increased her supplementation. She set a goal of using enough daily nutrition supplements to ensure weight maintenance. When discussing the use of supplements she said, "I don't want to spend

TABLE 19–1. Initial treatment goals, measures, and methods

Treatment goal	Measure	Method
Reduce depressive symptoms by 50% by 8 weeks	QIDS-SR_{16}	Pharmacotherapy Supportive therapy
Weight gain	Weekly weights	Supportive and behavioral therapy Pharmacotherapy Medical therapy
Reduce difficulty eating	Self-report	Supportive and behavioral therapy Pharmacotherapy Medical therapy
Reduce cancer trauma/stress	None	Supportive therapy Pharmacotherapy
Address relationship issues	Self-report, global	Supportive therapy, monitoring Couples therapy?

Note. QIDS-SR_{16}=Quick Inventory of Depressive Symptomatology–Self-Report.

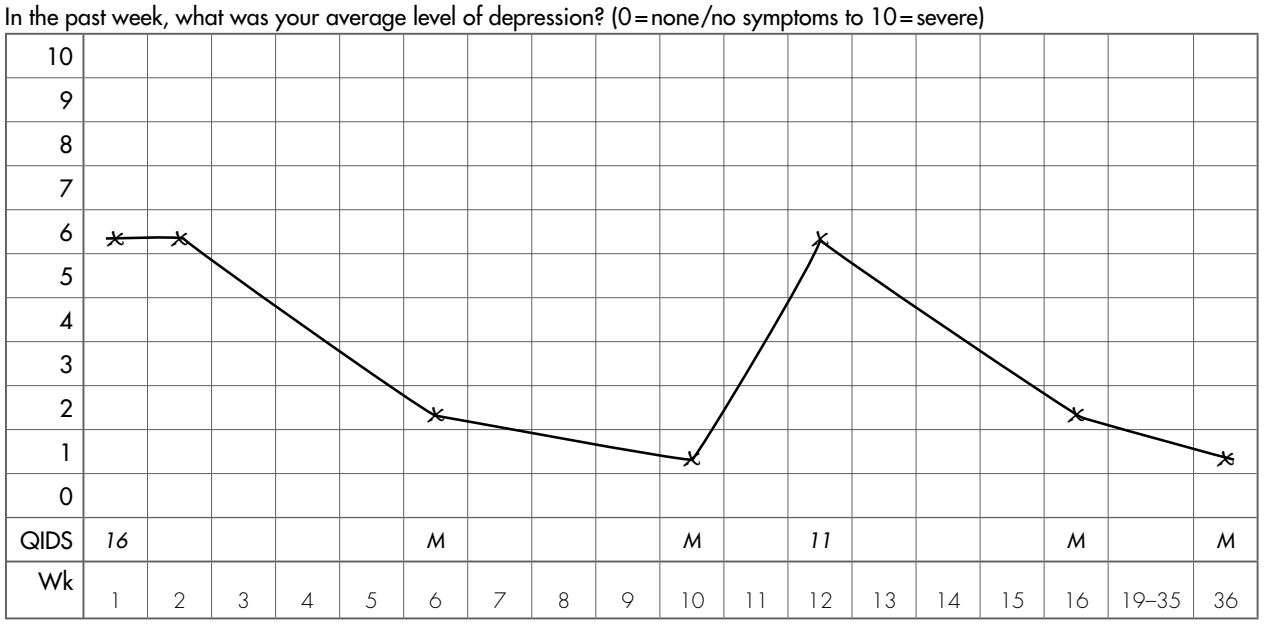


FIGURE 19–1. Depression trends (X–X).
M=missing; QIDS=Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}).

the rest of my life drinking this crap." We discussed alternative thoughts (e.g., "If I gain weight I will feel better," "I only have to do this for another few months," "If I can hold my weight steady I won't need a tube"). The patient was going to be out of town the following week, so the next visit was scheduled for 2 weeks.

Visit 4 (Week 5)

The patient's mood had improved and intake was up until 2 days before the visit when a magnetic resonance imaging scan found an "abnormality" in her colon. She was scheduled for a colonoscopy (and would need to be NPO [nothing by mouth]) and was terrified that her cancer had metastasized or that she might have some other medical problem. We discussed this new event as being a major trauma. From her food records, she had found it useful to think of her eating as, "This is what it's going to be like. If I have the odd good moment, yippee." We scheduled a visit after her colonoscopy.

Visit 5 (Week 6)

The colonoscopy was negative and the patient's mood was demonstrably improved. However, she continued to lose weight. We reviewed her food records, looking for strategies to improve eating and examining negative and catastrophic cognitions ("I hate this food. I will always hate this food"), as well as identifying activities that might give her some pleasure or fun or even relatively more pleasure or fun.

Visits 6–10 (Weeks 7–11)

(Not all visits were graphed.) The patient's mood continued to improve, and she was able to maintain her weight at 109–110 lb.

Visit 11 (Week 12)

The patient developed an infection at the back of her tongue which made eating almost impossible. Her weight dropped by another 2 lb. She also said that her husband told her that their relationship was unsatisfactory. Her mood worsened. She scored an 11 on the QIDS-SR_{16}. It would have been appropriate to consider other antidepressants, but her worse mood seemed related to her infection and her issues with her husband. When I started to focus on the food records and her weight, she became annoyed and insisted, "I need to talk about my marriage." I dropped my agenda and listened. She was very tearful in talking about how frightened she was that her marriage might not last or that her husband was frustrated with her progress. I asked if she wanted him to come to a session or if marital therapy might

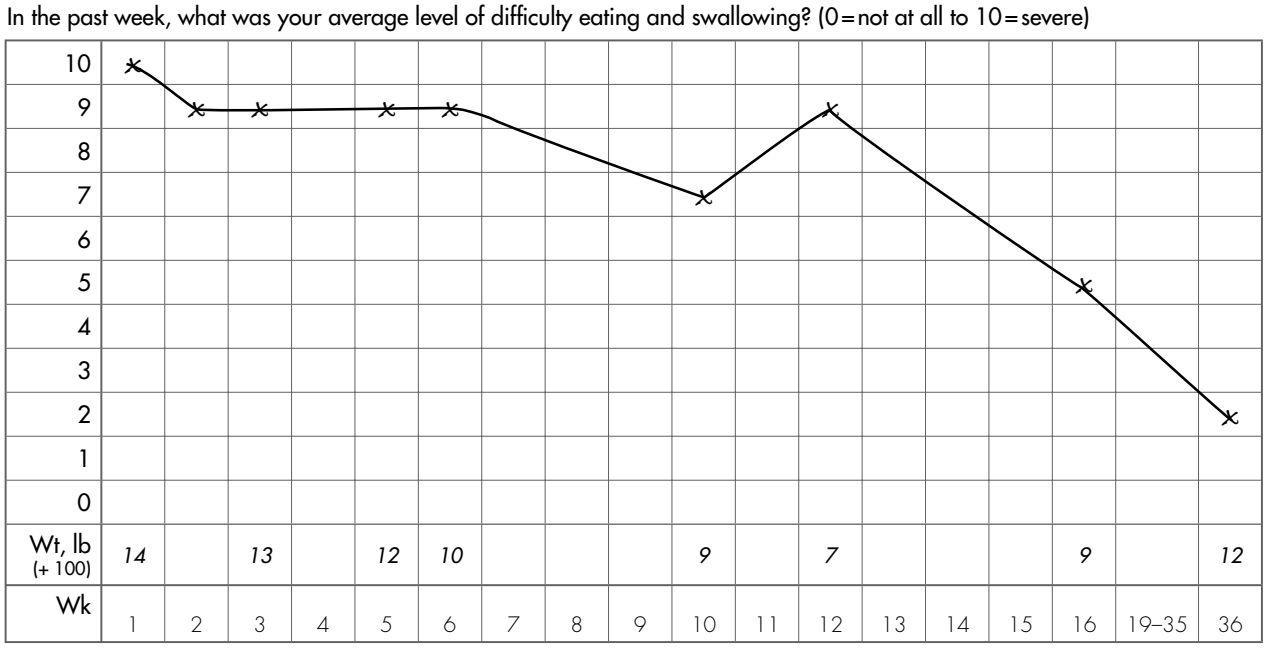


FIGURE 19–2. Eating/swallowing severity.

be appropriate. She said that she felt that their relationship was strong and had weathered other problems. A friend had invited them to spend some time with them in the desert. "If things aren't better when we get back, maybe we could see someone then." I *gently* encouraged her to continue her food records and weight monitoring while she was gone and thanked her for being very assertive about what she needed to talk about.

Visit 12 (Week 16)

The patient returned from the desert in a much better mood. Her QIDS-SR_{16} was now at 8, and her weekly average mood was a 2/10. She said that she had long talks with her husband and it seemed to help that he could "vent." They committed to continuing to work on their relationship. Her weight was 109 lb, which she felt was acceptable given all that had happened.

Final Visits

The patient continued to improve over the next several months. Her swallowing improved dramatically and she was able to eat regular, although soft, foods and to consume less Ensure. Taste remained a problem but she was optimistic; from what she had heard from online discussions and her support group, that would improve. Mood was normal. We scheduled a follow-up session in 3 months.

Summary of Guideline Use

I followed the general APA recommendations for major depressive disorder. I did not follow a strict CBT approach (use of thought records, etc.) but instead focused on the patient's need to maintain weight in the face of extreme discomfort.

I deviated from the STAR*D guideline as discussed above in minor ways for initial choice of medication. The patient had also "relapsed" by week 12 and, strictly on the basis of QIDS-SR, other interventions should have been considered. However, given the patient's exacerbated medical condition and relationship problems, it did not seem advisable to make a change until I had a sense of what would happen to her mood, assuming these problems would improve.

Sometimes, I follow the guidelines and manualized procedures too closely. If the patient had not interrupted me at week 12 to demand that we talk about what was really on her mind, I might not have been able to help her, could have overlooked an important contributing factor to her progress, and could have seriously impaired our therapeutic relationship. It can be challenging to listen carefully to

patients' concerns while trying to adhere strictly to structured interventions.

Ways to Improve Practice

In reviewing my notes and the guidelines (not reflected in the brief case discussion above), I realized that I was not carefully monitoring some aspects of the patient's care, especially in monitoring the use of Chinese herbs and alternative treatments. I also was adhering too rigidly to my treatment program at the expense of being insensitive to the patient.

References

Spiegel D, Classen C: Group Therapy for Cancer Patients: A Research-Based Handbook of Psychosocial Care. New York, Basic Books, 2000

Winston A, Rosenthal RN, Pinsker H: Introduction to Supportive Psychotherapy. (Core Competencies in Psychotherapy Series; Gabbard GO, series ed.). Washington, DC, American Psychiatric Publishing, 2004

Winzelberg AJ, Classen C, Alpers GW, et al: Evaluation of an Internet support group for women with primary breast cancer. Cancer 7:1164–1173, 2003

20

Anorexia, Obsessive-Compulsive Disorder, and Depression

C. Barr Taylor, M.D.

Introduction and Setting

See Chapter 17.

Illustration

This study illustrates:

- Use of obsessive-compulsive disorder (OCD) and anorexia guidelines
- Adaptation of guidelines and practices for a patient with limited financial resources
- Consideration of a new guideline as a new problem emerges

Chief Complaint

Ms. C.J. is a 35-year-old married woman who is self-referred for treatment of obsessive thoughts and anorexia.

Present Illness

Ms. C.J. said that about 4 months prior to admission (PTA), she began to have sexual fantasies about someone she worked with and also imagined that he had fallen in love with her. She was terrified that she was having these thoughts and wrote an e-mail to the man, indicating that she was not interested. Her husband was looking at her e-mail and discovered the note. He became very upset and threatened to divorce her. She told her husband that there was "nothing going on" and she would quit her job, which she did, to avoid further contact. However, she continued to obsess about her coworker. "I just can't get him out of my mind." She also began to feel that there were "germs" in her food. The thoughts seem to be getting worse and made her feel hopeless and depressed.

On the Yale-Brown Obsessive Compulsive Scale (Goodman et al. 1989a, 1989b), she reported current contamination, sexual, and somatic obsessions and fears, and hand-washing compulsions—she reported washing her hands about 20 times/day.

In addition, she reported a long history of food restriction and one episode of being hospitalized for anorexia about 5 years ago. About 2 months prior to assessment she began to reduce her food intake, but she said it was because of loss of appetite and not because of fear of germs. Over the previous 2 months, her weight had dropped from about 115 to 100 lb. She denied laxative use, purging, or driven exercise and said she wanted to weigh more.

One week prior to assessment her internist started her on sertraline 50 mg for her depression and she said that it helped but made her very "sluggish."

Other Significant Findings From Assessment

The patient reported no medical problems. She denied any alcohol or drug use. She reported multiple family members with depression. Her sister took fluoxetine with "good benefit." The patient was married for 12 years, with 2 children ages 3 and 11.

She was a high school dropout. On the mental status examination, her affect was very guarded and somewhat flat. There was no evidence of a thought disorder, and she denied hallucinations or delusions. Her baseline Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}) was 5, in the normal range. This rating seemed inconsistent with her affect, which seemed sad. She reported a baseline level of 7/10 for depression, 6/10 for anxiety, intrusive thoughts a few times per week, with the thoughts being 7/10 in terms of severity (on a 0- to 10-point rating scale, with 0=no symptoms, 10=severe symptoms). Her weight was 100 lb.

Differential Diagnosis

The patient met DSM-IV-TR (American Psychiatric Association 2000) criteria for OCD. She did not meet criteria for a current depressive episode, but her answers were inconsistent as she reported significant depression on a one-item scale but few symptoms on the QIDS-SR_{16} and attributed her 15-lb recent weight loss to a loss of appetite. She had a history of anorexia but denied current symptoms. However, she was very guarded on the initial interview and seemed to minimize symptoms. She was very uncomfortable in talking about her obsessive thoughts.

DSM-IV-TR Diagnosis

Axis I	OCD
	Probable current anorexia nervosa
	Major depressive episode?
Axis II	None
Axis III	None
Axis IV	Deferred
Axis V	GAF score: 50

Treatment Plan Considerations

The patient's main concerns related to her obsessive thoughts. She felt that her depressive symptoms were secondary to these thoughts and that she "desperately" wanted them to stop. The family's limited financial resources would be a factor in determining treatment. The patient lost her insurance coverage when she quit her job. Her husband's insurance had limited psychiatric benefits, and the family's income was just above the poverty line. A major focus of the initial sessions would be to build a sense of trust. I had the impression that she was afraid of revealing "too much."

Selecting a Guideline and Outcomes

The American Psychiatric Association (APA) OCD guideline (www.psychiatryonline.com/pracGuide/pracGuidehome.aspx) seemed to be an appropriate guideline to follow, at least initially. The selection of the guideline was complicated by the possibility of the patient having an eating disorder. The OCD treatment guideline states that treatment outcome targets include <1 hour/day spent obsessing and performing compulsive behaviors, no more than mild OCD-related anxiety, an ability to live with OCD-associated uncertainty, and little or no interference of OCD with the tasks of daily living. I decided to focus on a simple measure of frequency ("In the past week, how often did you experience upsetting thoughts?" rated from "never" to "a few times a day," and "In the last week, how severe was this symptom?" rated from "not bad at all" to "severe").

Anorexia

The APA eating disorders guideline would be reasonable to consider. Initially I decide to focus on weight gain and monitor how she was doing and to get a better sense of her attitudes and behaviors related to eating.

Depression and Anxiety Symptoms

The patient did not meet criteria for a depressive disorder but I was struck by her weight loss, which might have been due to depressive symptoms or anorectic food refusal (she denied the latter). Her score on the QIDS-SR_{16} was in the normal range, but I worried that she was minimizing symptoms. I decided to monitor depression and anxiety weekly and to entertain a focus on either depending on her course. Medication for her OCD would likely help her depression as well.

Treatment Plan and Initial Treatment

The OCD guideline states that the first-line treatments for OCD are cognitive-behavioral therapy (CBT) and serotonin reuptake inhibitors (SRIs). The guideline recommends that CBT focus on exposure and response prevention but notes the patient must be agreeable to and able to use the techniques. Having provided CBT to a number of

OCD patients, I was concerned that the patient would not have time to learn the techniques and that we would also need our limited number of sessions to deal with her eating and other problems. The anorexia guideline recommends fluoxetine (at a higher than usual dose). Because the patient was having unpleasant side effects on the sertraline, I decided to switch to fluoxetine. Other considerations in choosing fluoxetine were that her sister had done well on it and it is less expensive in case she needed to purchase the medication herself.

The OCD guideline recommends a course of CBT. However, the patient had a high school education and limited reading and writing skills, and her insurance coverage was limited to 10 sessions/year. The focus on therapy would need to be on medication and supportive therapy. I also mentioned that I would be weighing her before each session. I decided I would include CBT techniques as appropriate and as I had time (see Table 20–1).

Course

Visits 2–3 (Weeks 2 and 3)

The patient reported doing a little better in terms of improved mood, but her OCD thoughts remained unchanged. I printed out the National Institute of Mental Health handout on OCD to give her some background on the illness. The sessions focused on obtaining more information but also on providing some encouragement to eat. The patient felt that her continued obsessive thoughts about her former coworker were driving a wedge between her and her husband, and she wanted to find ways to spend more time with her husband. I also gave her a relaxation tape to help reduce her anxiety and increased the fluoxetine from 20 mg to 40 mg. Figures 20–1 and 20–2 show the course of her depressive and OCD symptoms and her weight.

Visit 4 (Week 6)

The patient reported that her mood had improved. She also felt some decrease in the frequency of her obsessive thoughts (not graphed in Figure 20–2) but no change in intensity of symptoms. She also reported significant problems eating; she was not able to increase her intake as requested—in fact, she wanted to further restrict. She did, however, report an improvement in her appetite. Her wish to restrict, history of anorexia, and thin-body ideal suggested that her weight loss was related to anorexia. We discussed the advantages of finding a support group for eating disorders. She continued to have difficulties with her husband. We discussed seeking a counselor, perhaps through her church.

The OCD guideline notes that it may take 8–12 weeks, including 4–6 weeks at a maximally tolerable dose, for the pharmacotherapy to be effective. The patient might have been able to tolerate a higher dose of fluoxetine, and using a higher dose would be consistent with the anorexia treatment guideline. However, as she had shown improvement, I decided not to increase the medication at this time. Adding CBT would be ideal but was not feasible for reasons mentioned above. I decided to start her on a low dose of risperidone (0.25 mg) to see if this might help her anxiety and help her gain weight. I gave the patient a self-monitoring food log to use at home. She was asked to simply note for each meal whether or not she ate "too little" or "enough."

TABLE 20–1. Initial treatment goals, measures, and methods

Treatment goal	Measure	Method
Reduce obsessive-compulsive disorder symptoms	Self-report frequency and severity: 50% reduction in frequency and severity in 6 months	Pharmacotherapy Supportive therapy Cognitive-behavioral therapy?
Weight gain	Weekly weights: target weight 110 lb in 6 months	Supportive and behavioral therapy
Reduce depressive symptoms	Self-report: 50% improvement by 8 weeks	Pharmacotherapy Supportive and behavioral therapy

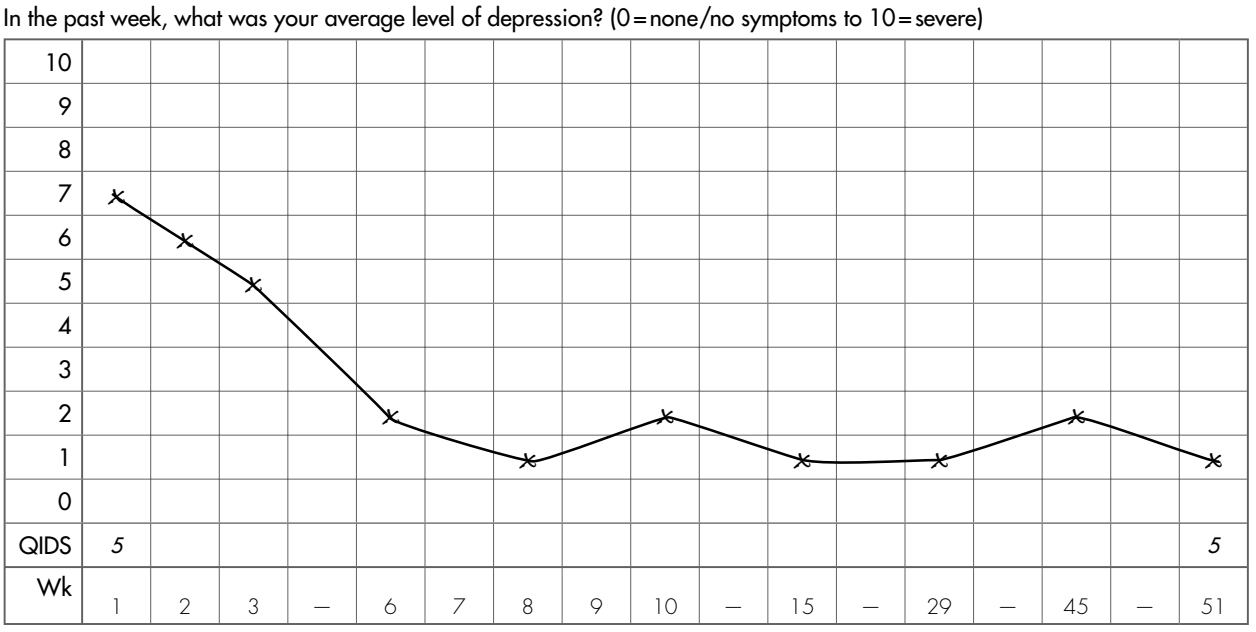


FIGURE 20–1. Depression trends (X–X).
QIDS=Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}).

Visit 5 (Week 8)

Ms. C.J. reported that her mood was much better and her anxiety much less. She had gained 2 lb. She discussed wanting to go back to work.

Visit 6 (Week 10)

The patient reported that her obsessive thoughts had gotten much worse. On careful questioning she reported that she had been "hearing voices." She felt that her former coworker had actually been talking to her. "I can hear his voice in my head." She said this had happened to her before she first came to see me, but she was afraid to tell me for fear of being "locked up." I reviewed my diagnosis and treatment plan. She reported a slight increase in her depression but did not meet criteria for major depressive episode with psychotic features. She did, however, meet the criteria for psychotic disorder not otherwise specified.

I reviewed the schizophrenia treatment guidelines. The APA practice guideline recommends the Brief Psychiatric Rating Scale (BPRS; Overall and Gorham 1962) and the Positive and Negative Syndrome Scale (PANSS; Kay et al. 1987) for monitoring psychopathology. The PANSS needs to be purchased and requires 30–40 minutes to administer and score (Kay et al. 1987), making it impractical in a clinical setting. On the BPRS, the patient had symptoms of anxiety, emotional withdrawal, guilt, hallucinatory behavior, and blunted affect for a total score of 19. A score of less than 18 is considered improved. I added the BPRS to her monitoring form (see Figure 20–2).

Although she had received some initial benefit from the risperidone, the dosage was below the therapeutic range used, for instance, in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) (Lieberman et al. 2005) or for augmenting treatment of OCD (Maina et al. 2008). Cost was a major factor. I discussed the cost/benefit of using perphenazine or haloperidol. She wanted to continue with the risperidone. I increased the risperidone to 0.5 mg/day and encouraged her to increase it over the next week to 2.0 mg/day.

Visit 7 (Week 15)

The patient reported a decrease in the voices, that her mood was slightly better, and that her OCD symptoms had returned to their level of a month or so ago. She said, however, that she had not increased the risperidone because she was not able to afford it. She had also gained 3 lb, which she was able to tolerate. Because of cost, she wanted to return in 3 months.

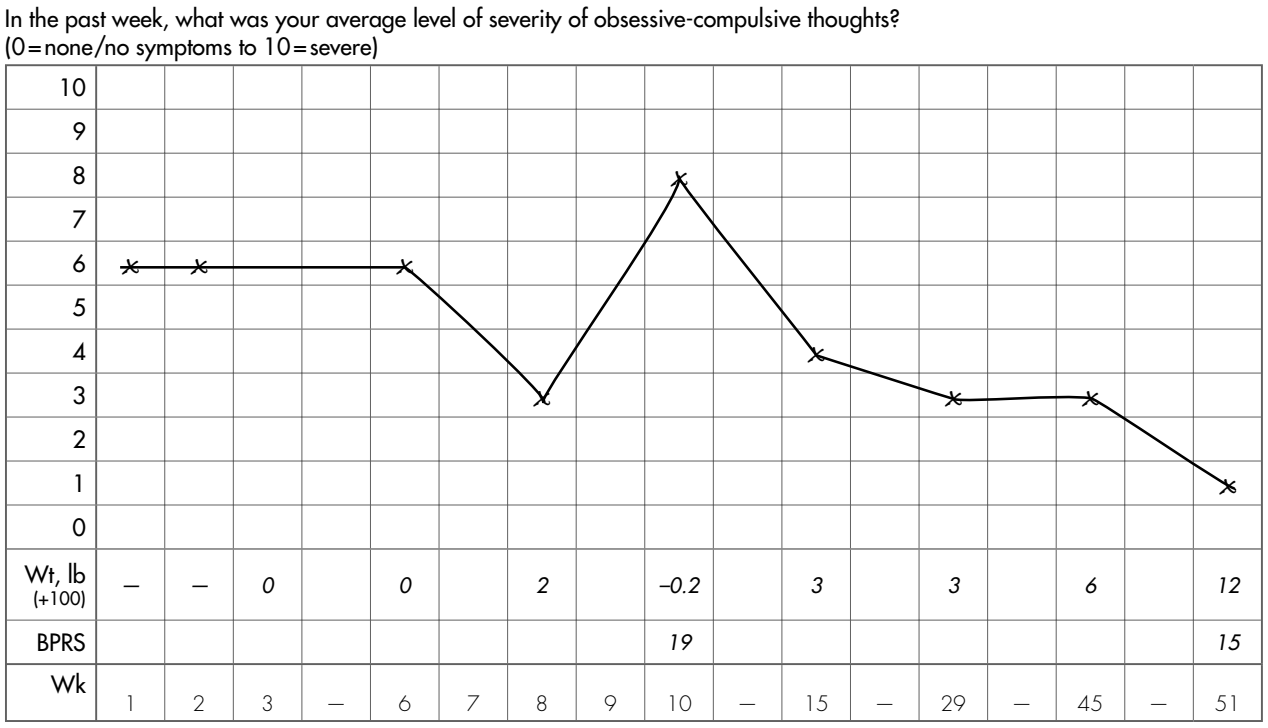


FIGURE 20–2. Severity of obsessive-compulsive thoughts (X–X).
BPRS=Brief Psychiatric Rating Scale.

Visit 8 (Week 29)

The patient reported that she had been doing OK. Her OCD symptoms were much less bothersome. She said she only heard occasional voices. She had maintained her 3-lb weight gain. I focused on issues related to her eating and weight and the need for her to continue taking her medication.

Visit 9 (Week 45)

A week after visit 8 the patient called to say that she stopped the risperidone because she had run out and couldn't afford it. After she had stopped the medication she still heard voices but she said they were less bothersome. She said that her "obsessive thoughts" (which she distinguished from the voices) were less. I prescribed a low dose of perphenazine. She had not attended a support group for eating disorders because she said that she had no one to leave her children with. We problem-solved ways for her to find support for her attending such a group.

Visit 10 (Week 51)

The patient never filled the perphenazine prescription. However, she continued to feel less depressed and anxious. She still had mild obsessive thoughts and heard occasional voices, but they continued to be less troublesome. She was also able to gain another 6 lb. The patient planned to move from the area. We discussed options for her continued care, including going to a community mental health center when she had settled into another area.

Summary of Guideline Use

This patient presented a number of challenges, including the need to address a number of complex issues in a limited number of sessions with a patient with limited financial resources and little support from her husband. Four different guidelines provided useful guidance for her treatment. I began by following the APA OCD guideline. As seen in Table 20–2, I followed the General Psychiatric Management principles (A1–8). For the initial treatment, CBT was recommended, but the patient was not able to afford the required number of sessions and didn't seem to be a good candidate for self-help. (I have not included my checklist for the rest of the guidelines.) At week 6, I decided to augment with an atypical antipsychotic because she was still having

TABLE 20–2. American Psychiatric Association general recommendations for obsessive-compulsive disorder

	F	C
A. Psychiatric management		
1. Establish and maintain a therapeutic alliance		
Tailor communication style to the patient's needs and abilities	√	
Allow patients with excessive worry or doubting time to consider treatment decisions; repeat explanations if necessary	NA	
Attend to transference and countertransference, which may disrupt the alliance and adherence	√	
Consider how the patient's expectations are affected by his or her cultural and religious background, beliefs about the illness, and experience with past treatments	√	
2. Assess the patient's symptoms		
Use DSM-IV-TR criteria for diagnosis	√	
Consider using screening questions to detect commonly unrecognized symptoms	√	
Differentiate OCD obsessions, compulsions, and rituals from similar symptoms found in other disorders	√	
3. Consider rating the patient's symptom severity and level of function. The following instruments are useful:		
For symptoms: Yale-Brown Obsessive Compulsive Scale	√	
For self-rated depression: Patient Health Questionnaire (PHQ-9), Beck Depression Inventory–II, Zung Depression Scale, and the patient versions of the Inventory of Depressive Symptoms (IDS) or Quick Inventory of Depressive Symptomatology (QIDS)	√	
For disability: Sheehan Disability Scale	No	
For quality of life: Quality of Life Enjoyment and Satisfaction Questionnaire or the more detailed World Health Organization Quality of Life Survey	No	
4. Enhance the safety of the patient and others		
Assess for risk of suicide, self-injurious behavior, and harm to others	√	
Consider obtaining collateral information from family members and others		√
Take into consideration factors associated with increased risk of suicide, including specific psychiatric symptoms and disorders and previous suicide attempts		√
Evaluate the patient's potential for harming others, either directly or indirectly (e.g., when OCD symptoms interfere with parenting)		√
5. Complete the psychiatric assessment		
See APA's (1995) *Practice Guideline for the Psychiatric Evaluation of Adults*	√	
Assess for common co-occurring disorders, including mood disorders, other anxiety disorders, eating disorders, substance use disorders, and personality disorders	√	
6. Establish goals for treatment; goals of treatment include decreasing symptom frequency and severity, improving the patient's functioning, and helping the patient to improve his or her quality of life	√	
Reasonable treatment outcome targets include < 1 hour/day spent obsessing and performing compulsive behaviors, no more than mild OCD-related anxiety, an ability to live with OCD-associated uncertainty, and little or no interference of OCD with the tasks of daily living; despite best efforts, some patients will be unable to reach targets	√	

TABLE 20–2. American Psychiatric Association general recommendations for obsessive-compulsive disorder *(continued)*

	F	C
7. Establish the appropriate setting for treatment		√
Generally choose least restrictive setting that is safe to allow for effective treatment		√
Is hospitalization indicated?		√
Is home-based treatment needed?		√
8. Enhance treatment adherence		
Recognize that the patient's fears, doubting, and need for certainty can influence his or her willingness and ability to cooperate with treatment and can challenge the clinician's patience	√	
Provide education about the illness and its treatment, including expected outcomes and time and effort required	±	
Inform the patient about likely side effects (SEs), inquire about SEs the patient may be unwilling to report (e.g., sexual SEs), respond quickly to concerns about SEs, and schedule follow-up appointments soon after starting or changing medications	√	
Address breakdowns in the therapeutic alliance	NA	
Consider the role of the patient's family and social support system		√
When possible, help the patient to address practical issues such as treatment cost, insurance coverage, and transportation	√	

Note. The APA guidelines contain both recommended actions and areas for consideration. As in the author's example above, you may use the columns to the right to indicate that a recommendation was followed (F) or considered (C), as appropriate. You may also indicate NA for "not applicable."
APA=American Psychiatric Association; OCD=obsessive-compulsive disorder.

Source. Adapted from http://www.psychiatryonline.com/pracGuide/pracGuidehome.aspx.

disabling symptoms and extensive anxiety. Following the OCD guidelines, I might have increased the fluoxetine and waited several more weeks before trying an atypical antipsychotic.

I also considered the APA general eating disorder guidelines and those specific to anorexia (see Table 20–3 for those related to anorexia). The general guideline encourages the therapist to address a number of "underlying themes and correct core maladaptive thoughts and attitudes" (B.6). I did not have time to address many of these issues. Unfortunately, there is no strong evidence to guide treatment of adult anorexia patients (Keel and Haedt 2008). I might have done a better job in providing some simple psychoeducational materials. Overall, I don't think I provided enough homework activity; more might have served as a bridge between sessions.

The APA schizophrenia guideline was also relevant. This was not a major focus of therapy as the patient's "auditory hallucinations" seems to respond to a low dose of risperidone, and in fact they remained improved even when she no longer took the medication. The APA depression guideline also provided useful and different information on how to manage this patient. Because the use of this guideline is presented in Chapter 18, I don't go over it here. Overall, I ended up selecting pieces of the guidelines, perhaps at the expense of not providing some important therapeutic procedures. Would I have done better, for instance, to focus on only one type of CBT practice, such as thought-stopping? Would this even have clarified the relationship between her thoughts and her "hallucinations"?

Ways to Improve Practice

The patient did much better than I expected. As presented, the use of the fluoxetine would seem to have been a major factor in her improvement, but I suspect other nonspecific elements may have been important, such as providing a supportive and trusting relationship, which allowed her to discuss a number

TABLE 20–3. Specific recommendations for anorexia adapted from American Psychiatric Association eating disorders guideline

		F
a.	Nutritional rehabilitation	
	Establish goals for seriously underweight patients	√
	Restore weight	±
	Normalize eating patterns	√
	Achieve normal perceptions of hunger and satiety	No
	Correct biological and psychological sequelae of malnutrition	NA
	Help the patient to resume eating and gain weight (see guidelines for details)	√
	Help the patient to maintain weight	NA
b.	Psychosocial treatments	
	Establish goals, including to help the patient	√
	Understand and cooperate with nutritional and physical rehabilitation	√
	Understand and change the behaviors and dysfunctional attitudes related to the eating disorder	√
	Improve interpersonal and social functioning	√
	Address comorbid psychopathology and psychological conflicts that reinforce or maintain eating disorder behaviors	±
	Establish and maintain a psychotherapeutically informed relationship with the patient	√
	Provide formal psychotherapy once weight gain has started.	No
c.	Medications	
	Use psychotropic medications in conjunction with psychosocial interventions, not as a sole or primary treatment for patients with anorexia nervosa	√
	Whenever possible, defer making decisions about medications until after weight has been restored	NA
	Be aware of and manage general side effects	√
	Consider antidepressants to treat persistent depression or anxiety following weight restoration	√
	Consider approaches to restore lost bone mineral density	No

Note. F=followed; NA=not applicable.

Source. http://www.psychiatryonline.com/pracGuide/pracGuidehome.aspx.

of distressing issues. In reviewing the case and my use of guidelines, here are two ways in which I could improve my practice: 1) use higher doses for longer periods of the initial SRI in treating patients with OCD, and 2) find low-literacy, simple psychoeducational materials to supplement therapy. Given the unusual picture and complexity of this patient's -psychopathology, a consult with a peer would have been warranted.

References

American Psychiatric Association: Practice guideline for the psychiatric evaluation of adults, 2nd edition, in Practice Guidelines for the Treatment of Psychiatric Disorders, Compendium. Washington, DC, American Psychiatric Publishing, 2006, pp 1–60

Goodman WK, Price LH, Rasmussen SA, et al: The Yale-Brown Obsessive Compulsive Scale, I: development, use, and reliability. Arch Gen Psychiatry 46:1006–1011, 1989a

Goodman WK, Price LH, Rasmussen SA, et al: The Yale-Brown Obsessive Compulsive Scale, II: validity. Arch Gen Psychiatry 46:1012–1016, 1989b

Kay SR, Fiszbein A, Opler LA: The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull 13:261–276, 1987

Keel PK, Haedt A: Evidence-based psychosocial treatments for eating problems and eating disorders. J Clin Child Adolesc Psychol 37:39–61, 2008

Lieberman JA, Stroup TS, McEvoy JP, et al: Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) investigators: effectiveness of antipsychotic drugs in patients with chronic schizophrenia. N Engl J Med 353:1209–1223, 2005

Maina G, Pessina E, Albert U, et al: 8-week, single-blind, randomized trial comparing risperidone versus olanzapine augmentation of serotonin reuptake inhibitors in treatment-resistant obsessive-compulsive disorder. Eur Neuropsychopharmacol 18:364–372, 2008

Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. Psychol Rep 10:799–812, 1962

This page intentionally left blank

21

Major Depressive Disorder

A Measurement-Based Care Approach

Madhukar H. Trivedi, M.D.
Ben Kurian, M.D., M.P.H.

Setting

The following case is based on a patient who participated in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D), a research trial using measurement-based care (MBC) to treat major depressive disorder (MDD) in real-world outpatients with depression.

Illustration

This study illustrates:

- Use of a nonpsychotic major depressive episode sequential treatment algorithm
- Use of STAR*D MBC guidelines
- Personalized treatment of MDD and disease self-management

Chief Complaint

Ms. S.W. is a 40-year-old married white female who works as a truck driver and presents with symptoms of depression.

Present Illness

Ms. S.W. stated that she had been depressed for the last 6 months, following the death of her mother, and the depression had gotten particularly worse in the past 2 months. The patient stated that her mother had a history of high blood pressure and high cholesterol and had died suddenly from a heart attack 6 months prior to the patient's initial clinic presentation. For the past 2 months the patient felt sad nearly every day and had decreased sleep (initial insomnia), decreased interest in her usual activities, increasing fatigue, and feelings of worthlessness and guilt. In addition, the patient reported increasing difficulties with concentration but denied change in appetite or weight. She also denied any thoughts of suicide or preoccupation with death. Her symptoms had worsened to the point at which they interfered with her ability to concentrate while at work, with the patient stating that in the past 2 months she had gotten lost while driving (which was very unusual for her). Ms. S.W. had remarried recently (about 3 months ago), although she denied any significant stress associated with this event.

Past Psychiatric History

The patient reported one prior major depressive episode, approximately 10 years ago while undergoing a divorce from her first husband. At that time, the patient noted her primary care physician started her on sertraline and titrated her dose up to 100 mg/day,

which she took for 1 year with good results. She stopped the sertraline on her own because she felt she "no longer needed it." In addition to sertraline, the patient stated her primary care physician also prescribed Ambien (zolpidem) to assist with initial insomnia.

Medical History

- Hypertension (stable)—taking hydrochlorothiazide 25 mg
- Gastroesophageal reflux disease (stable)—taking pantoprazole 40 mg
- Tension headaches (stable)—taking ibuprofen as needed
- Hysterectomy 16 years ago—not taking hormone replacement

Social History

The patient denied tobacco use. She drank approximately two to three cups of coffee daily. She reported no routine exercise. She stated that she rarely had alcohol (approximately one glass of wine per month). She denied use of drugs or other pharmaceuticals (i.e., herbal medications). The patient lived with her husband and his teenage son. She had been living with her current husband for some time before they got married. She has one prior marriage with two now-adult children. The patient graduated from high school but reported no further schooling.

There was no significant family history of depression or mental illness. The patient's mental status examination was significant for a depressed affect with limited range of emotion. Her thought process and content were intact, and her Mini-Mental State Examination score = 30; however, the patient subjectively reported distress with short-term memory and concentration. Her physical exam and vital signs were within normal limits. Her baseline Quick Inventory of Depressive Symptomatology–Clinician Rated (QIDS-C_{16}) score was 19, in the severe range (see below for interpretation of scores):

QIDS-C_{16} scoring criteria:

Normal	≤5
Mild	6–10
Moderate	11–15
Severe	16–20
Very severe	21–27

DSM-IV-TR Diagnosis

Axis I	Major depressive disorder, recurrent, current episode severe, without psychotic features
Axis II	None
Axis III	Hypertension, gastroesophageal reflux disease, tension headaches
Axis IV	Recent loss of mother
Axis V	GAF score: 55

Treatment Plan Considerations

Major Depressive Disorder and Uncomplicated Bereavement

The patient's main problem was her depression. She was concerned that if it continued, it would affect her new marriage. She stated that the grief surrounding her mother's sudden death had improved a few months previously but that she was still depressed. The patient felt guilty that she remained depressed despite recently getting remarried. She had not had any counseling/psychotherapy regarding her recent loss and wished to first try pharmacological treatment for MDD.

MDD Treatment Guideline Selection

A number of treatment algorithms and guidelines are available to clinicians considering the treatment of MDD: 1) the STAR*D guideline (Rush et al. 2004), 2) the American Psychiatric Association (2000) guidelines for MDD, and 3) the TMAP guideline for nonpsychotic MDD (Crismon et al. 1999).

The goal of treatment for all patients with MDD, as outlined in Rush et al. (2006b), is to achieve remission of depressive symptoms (i.e., the patient is symptom free). The TMAP guideline for MDD introduced the concept of using critical decision points (CDPs) at defined intervals to objectively assess symptom and side-effect burden and to adjust treatment accordingly (Crismon et al. 1999). STAR*D further refined the objective assessments of symptoms and side effects with the introduction of measurement-based care (Trivedi and Daly 2007; Trivedi et al. 2006b, 2007b).

MBC promotes the use of rating scales or questionnaires to measure symptoms, side effects, and adherence at every visit as well as to guide tactics to modify dosage and treatment duration. In addition, MBC integrates these objective measurements and

targets for full symptom remission. Most clinicians do not routinely use specific measures of depressive symptoms at patient visits; rather, they tend to use global measures (Biggs et al. 2000). Results from level 1 of STAR*D reveal that the use of MBC may lead to greater remission rates than those seen in efficacy studies for patients with chronic depression (Trivedi et al. 2006b).

STAR*D used the QIDS, a 16-item questionnaire based on DSM-IV-TR criteria for MDD, developed as a measure of depressive symptom severity (www.ids-qids.org). The QIDS is quick and easy to use, with good reliability and validity. Both clinician and self-rated versions are available free of charge and have been validated and translated into other languages (Rush et al. 2003, 2006a; Trivedi et al. 2004). To comply with the STAR*D MBC guidelines, we present the QIDS in this case presentation, but we recognize that alternative symptom ratings such as the Patient Health Questionnaire–9 (PHQ-9; Kroenke et al. 2001), Beck Depression Inventory–II (Steer et al. 1999), and others can be used. With regard to systematically measuring side effects, MBC guidelines recommend using the Frequency, Intensity, and Burden of Side Effects Rating (FIBSER; Wisniewski et al. 2006).

Table 21–1 reflects MBC for the acute treatment phase of MDD. Clinical status (e.g., remission, partial response) and side-effect tolerability at each CDP determine the treatment plan for that visit. Remission was defined by QIDS-C_{16}≤5, partial response was defined as QIDS-C_{16}=6–8, and nonresponse essentially was defined as QIDS-C_{16}≥9. Clinic visits in STAR*D were generally recommended at weeks 0, 2, 4, 6, 9, and 12 at each treatment stage (level) or until an adequate remission or response was obtained, with CDPs occurring at each of these visits (except week 2).

TABLE 21–1. Critical decision points (CDPs) and tactics for acute phase treatment of major depression

Critical decision point	Clinical status	Plan
Week 0 (CDP #1)	Symptomatic	Initiate medication; adjust dose to lower end of therapeutic dose range or serum level
Week 4 (CDP #2)	Remission (QIDS-C_{16}≤5)	Continue current dose
	Partial response (QIDS-C_{16}=6–8)	Continue current dose Consider increasing dose
	Nonresponse (QIDS-C_{16}≥9)	Increase dose
Week 6 (CDP #3)	Remission (QIDS-C_{16}≤5)	Continue current dose
	Partial response (QIDS-C_{16}=6–8)	Increase/maximize dose
	Nonresponse (QIDS-C_{16}≥9)	
Week 9 (CDP #4)	Remission (QIDS-C_{16}≤5)	Continue current dose
	Partial response (QIDS-C_{16}=6–8)	Increase dose Go to the next level
	Nonresponse (QIDS-C_{16}≥9)	Discontinue and go to the next level
Week 12 (CDP #5)	Remission (QIDS-C_{16}≤5)	Go to follow-up phase
	Partial response (QIDS-C_{16}=6–8)	Go to the next level Increase dose and reevaluate in 2 weeks

Note. QIDS-C_{16}=Quick Inventory of Depressive Symptomatology–Clinician Rated.

Course

Level 1 Treatment Course (Citalopram)

CDP 1 (Week 0)

Prior to entering a patient into any treatment algorithm, one should conduct a thorough diagnostic evaluation to confirm a primary diagnosis of MDD and to uncover any other potential comorbid medical and psychiatric conditions. Antidepressant selection should be based on individual patient characteristics (e.g., history of prior treatment response) and patient preference. Additional consideration for side-effect profiles, potential drug interactions, and comorbid medical and psychiatric conditions may favor the choice of one particular medication over another.

In this case example, the patient, diagnosed with MDD without psychotic features (QIDS-C_{16}=19), was started on citalopram 20 mg. However, at this treatment stage, it is acceptable to prescribe any of the U.S. Food and Drug Administration (FDA)–approved selective serotonin reuptake inhibitors (SSRIs; e.g., fluoxetine, paroxetine, sertraline, citalopram, and escitalopram); serotonin and norepinephrine reuptake inhibitors (SNRIs; e.g., venlafaxine XR and duloxetine), or bupropion SR or XL. While head-to-head comparisons are not available for all antidepressants, data from the FDA and comparison trials suggest that all these antidepressants are similar in efficacy in treatment-naïve patients (level A data; Hansen et al. 2005).

The patient was counseled on potential medication side effects and the potential increased risk of suicidal thinking and behavior (i.e., suicidality), in accordance with FDA guidelines (U.S. Food and Drug Administration, Center for Drug Evaluation and Research 2004). Additionally, prior to starting a patient on an antidepressant, both the patient and the patient's family should be educated regarding the need to monitor for suicidality and associated symptoms, including anxiety, agitation, insomnia, irritability, hostility, aggressiveness, impulsivity, akathisia (psychomotor restlessness), hypomania, mania, or any other unusual changes in behavior. These symptoms, if present, are likely to emerge early in the course of treatment, and therefore a return visit was scheduled for this patient in 2 weeks. Furthermore, a telephone follow-up was scheduled with the patient 1 week after initiating citalopram to assess tolerability and suicidality.

The following figures reflect the practical application of the STAR*D level 1 algorithm for the present case, at each of the critical decision points, beginning with Figure 21–1, CDP week 0.

Week 1

A phone screen was conducted between the patient and a research assistant 1 week after initiating citalopram 20 mg to assess tolerability and adherence, and specifically to assess for any increase in suicidal thoughts and/or behavior. The patient reported mild gastrointestinal side effects (i.e., cramps, diarrhea); however, overall she tolerated citalopram 20 mg and reported no missed doses. Additionally, the patient denied any suicidal thoughts or associated symptoms and was reminded of her follow-up appointment in 1 week.

Week 2

At week 2 visit, once again side effects, tolerability, and adherence were measured. The patient reported resolution of gastrointestinal side effects and once again denied suicidal ideation. She did report two missed medication doses, when she forgot to take her medication with her while staying with her mother. The next visit, week 4, included a CDP to determine a treatment decision based on measurement of symptoms, side effects, and tolerability.

CDP 2 (Week 4)

According to the STAR*D MBC algorithm, the second CDP took place at week 4 (Figure 21–2). At this visit the patient had three domains of treatment measured: symptom severity (measured by the QIDS-C_{16}), side effects (measured by the FIBSER), and medication adherence. After 4 weeks of treatment with citalopram 20 mg, the patient's QIDS-C_{16} was measured at a 13. The patient noted slight improvements in mood and concentration; however, insomnia and anhedonia still persisted. The patient reported no side effects related to medication treat-

CDP, Week 0	STAR*D Level 1
	Start patient on citalopram 20 mg/day
Return to clinic:	Return in 2 weeks

FIGURE 21–1. Critical decision point (CDP) 1 (level 1).

CDP, Week 4	STAR*D Level 1
Symptom Improvement (SEs tolerable)	
QIDS-$C_{16} \geq 9$	Increase dose to 40 mg/day
QIDS-C_{16} = 6–8	Continue current dose, *or* Increase dose to 40 mg/day
QIDS-$C_{16} \leq 5$	Continue current dose
Improved, but SEs are intolerable	Continue current dose and address SEs, *or* Go to the next level
Not improved and SEs are intolerable	Go to the next level
Return to clinic:	Return in 2 weeks

FIGURE 21–2. Critical decision point 2 (level 1).
QIDS-C_{16}=Quick Inventory of Depressive Symptomatology; SEs=side effects.

ment. This was reflected by the FIBSER, which measured the frequency, intensity, and burden of side effects (the patient reported "no side effects" to all). Furthermore, the patient reported no additional missed doses of medication since the last visit. Specifically, because the patient's QIDS score remained≥9 and the side-effect profile was tolerable, the MBC algorithm dictated that the dose of citalopram be increased to 40 mg/day. Had the patient's side-effect profile for citalopram been intolerable, the algorithm does allow the clinician to go to the next level of treatment (i.e., switch to another antidepressant) at week 4 and beyond. Lastly, as was done at every visit, suicidal behavior and associated symptoms were assessed (patient reported none), and the patient was scheduled for follow-up visit in 2 weeks.

CDP 3 (Week 6)

At week 6 the patient scored a 14 on her QIDS-C_{16} and reported no significant improvement in depressive symptom severity. She also reported a new side effect she believed was induced by citalopram: daytime sleepiness (the patient had been taking citalopram daily in the morning), most notably present at work. According to the FIBSER, the patient reported daytime somnolence as present 25% of the time with mild intensity and impairment, but still tolerable. It was suggested to the patient to try and switch the time of citalopram dosing to evening to see if this improved side effect. Once again, according to the MBC algorithm (Figure 21–3), the daily dose of citalopram was increased to 60 mg, given acceptable side effects and persistence of depressive symptoms. Follow-up was scheduled in 3 weeks.

CDP 4 (Week 9)

The patient returned at week 9, after having taken citalopram 60 mg for 3 weeks. This visit was the first CDP that allowed for the clinician to move to the next treatment level due to an inadequate treatment response, despite tolerable side effects (Figure 21–4). At this visit the patient's depressive symptoms were virtually unchanged from the previous visit, QIDS-C_{16}=13, and only mildly improved from baseline. The side effect of daytime somnolence resolved, and the patient was uncertain whether this was secondary to switching the medication dosing time to evening or whether the increase to 60 mg daily played a role. Regardless, the patient reported no presence of side effects, with no intensity and no functional impairment on the FIBSER, and once again the patient also had not missed any medication doses in the past 3 weeks. However, with a QIDS≥9 after 9 weeks of treatment with citalopram, the clinician and patient elected at this CDP to move to the next level of treatment. According to the STAR*D algorithm, treatment options at level 2 are divided into switch and augmentation (Rush et al. 2004). The decision to switch or augment is determined by treatment response and medication tolerability. For example, if a patient has had no treatment response or has intol-

CDP, Week 6	STAR*D Level 1
Symptom Improvement (SEs tolerable):	
QIDS-$C_{16} \geq 9$	Increase dose to 60 mg/day
QIDS-$C_{16} = 6–8$	Increase dose to 60 mg/day, *or* Continue current dose
QIDS-$C_{16} \leq 5$	Continue current dose
Improved, but SEs are intolerable	Continue current dose and address SEs, *or* Decrease dose and continue for 2 additional weeks, *or* Go to the next level
Not improved and SEs are intolerable	Go to the next level
Return to clinic:	Return in 3 weeks

FIGURE 21–3. Critical decision point 3 (level 1).

erable side effects, a switch to another medication is recommended. According to STAR*D, switch options for level 2 include any of the following: sertraline, bupropion, venlafaxine, or cognitive therapy. However, if a patient has a tolerable side-effect profile and a partial treatment response, as is the case in this example, then a treatment augmentation is recommended. Level 2 treatment options for augmentation per the STAR*D algorithm include one of the following: buspirone, bupropion, or cognitive therapy (Rush et al. 2004). Figure 21–5 shows the patient's STAR*D level 2 treatment.

At the week 9 CDP visit, augmentation treatment options were discussed with the patient, and the only treatment she was unwilling to accept was cognitive therapy. According to STAR*D level 2 results, buspirone and bupropion are both efficacious as augmentation agents; however, bupropion is associated with better tolerability and greater reduction in self-reported depressive symptoms (Trivedi et al. 2006a). Therefore, bupropion XL (bupropion SR is also acceptable) was chosen as the level 2 augmentation agent, and the MBC algorithm for bupropion XL augmentation was initiated.

CDP, Week 9	STAR*D Level 1
Symptom Improvement (SEs tolerable):	
QIDS-$C_{16} \geq 9$	Go to the next level
QIDS-$C_{16} = 6–8$	Increase dose to 60 mg/day, if not done before, *or* Go to the next level
QIDS-$C_{16} \leq 5$	Continue current dose
SEs are intolerable	Go to the next level
Return to clinic:	Return in 3 weeks

FIGURE 21–4. Critical decision point 4 (level 1).

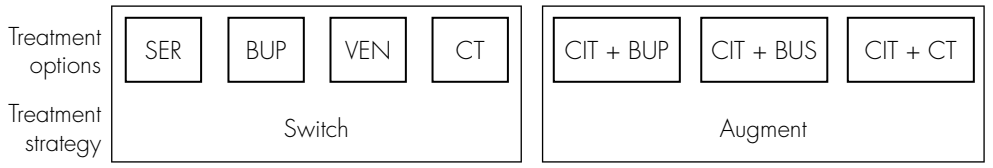


FIGURE 21–5. STAR*D level 2 treatment.

BUP=bupropion; BUS=buspirone; CIT=citalopram; CT=cognitive therapy; SER=sertraline; VEN=venlafaxine.

Source. Adapted from Rush AJ, Fava M, Wisniewski SR, et al: Sequenced Treatment Alternatives to Relieve Depression (STAR*D): rationale and design. *Controlled Clinical Trials* 25:119–142, 2004.

Level 2 Treatment Course (Citalopram + Bupropion XL)

CDP 1 (Week 0)

Prior to initiating any second treatment options, switch or augmentation, the clinician should reassess the diagnosis of MDD, with specific care to rule out a bipolar diathesis. For this case, after confirmation of diagnosis, the patient was continued on citalopram 60 mg, and bupropion XL 150 mg daily was initiated (Figure 21–6). In fact, during level 2 treatment, assuming tolerability it was recommended that the patient continue citalopram at the last acceptable dosage from level 1 (in this case 60 mg). Side-effect profile of bupropion XL and precautions to watch for increase in suicidal thoughts/behavior and associated symptoms were once again discussed with the patient. Follow-up was scheduled in 2 weeks.

Week 2

The patient returned for week 2 visit and underwent a routine clinic visit. She noted a mild headache the first week after bupropion was initiated with resolution of side effect in the days leading up to week 2 visit. She noted feeling better and denied any missed doses of either medication. Follow-up visit was scheduled in 2 weeks.

CDP 2 (Week 4)

Systematic measurement of symptoms, side effects, and adherence once again guided the level 2 CDP. At this visit, the patient noted significant improvement in depressive symptoms, with a QIDS-C_{16}=6, and specific improvements in mood, energy, concentration, and guilt. In fact, the only depressive symptom that the patient continued to complain of was anhedonia. The patient denied any continuing side effects (i.e., headache), and according to the FIBSER, side effects were determined to be tolerable. Additionally, the patient denied any suicidal thoughts/behavior or associated symptoms. At this stage, MBC provides the clinician with the choice of continuing on the current dose of the augmentation medication or increasing the dose if acceptable to the patient (Figure 21–7). The clinician and the patient jointly decided to increase the bupropion to 300 mg with a goal of targeting full symptom remission (QIDS-$C_{16}\leq 5$), and the patient was scheduled for follow-up in 2 weeks.

An emphasis has been placed on remission in recent years, and the American College of Neuropsychopharmacology Task Force advocates remission as the goal of treatment for patients with MDD (Rush et al. 2006b). This is largely due to the conse-

CDP, Week 0	STAR*D Level 2: Citalopram + Bupropion XL
	Start patient on bupropion 150 mg/day in addition to current dose of citalopram
Return to clinic:	Return in 2 weeks

FIGURE 21–6. Critical decision point 1 (level 2).

CDP, Week 4	STAR*D Level 2: Citalopram + Bupropion XL
Symptom Improvement (SEs tolerable)	
QIDS-$C_{16} \geq 9$	Increase dose to 300 mg/day
QIDS-C_{16} = 6–8	Continue current dose, *or* Increase dose to 300 mg/day
QIDS-$C_{16} \leq 5$	Continue current dose
Improved, but SEs are intolerable	Continue current dose and address SEs, *or* Go to the next level
Not improved and SEs are intolerable	Go to the next level
Return to clinic:	Return in 2 weeks

FIGURE 21–7. Critical decision point 2 (level 2).

quence of not achieving remission. Paykel and colleagues found that patients with residual depressive symptoms have an increased risk for and a shorter time to relapse of MDD (Paykel et al. 1995).

CDP 3 (Week 6)

The patient returned for her week 6 visit in high spirits. She noted feeling "so much better," and this was confirmed with a QIDS-C_{16}=3. Improvements were noted in all domains, and the patient officially achieved remission status. She was able to tolerate the 300 mg dose of bupropion XL without side effects, and the FIBSER again revealed no side-effect frequency, intensity, or burden. Furthermore, the patient was compliant with prescribed medications and reported no missed doses. Thus, the patient was continued on citalopram 60 mg and bupropion XL 300 mg (Figure 21–8), with a follow-up visit scheduled in 3 weeks.

CDP 4 (Week 9)

Results from week 9 visit mirrored the findings from week 6, in that the patient continued to remit and showed acceptable tolerability. Once again, the patient denied any suicidal thoughts/behavior or asso-

CDP, Week 6	STAR*D Level 2: Citalopram + Bupropion XL
Symptom Improvement (SEs tolerable)	
QIDS-$C_{16} \geq 9$	Increase dose to 450 mg/day.
QIDS-C_{16} = 6–8	Increase dose to 450 mg/day, *or* Continue current dose.
QIDS-$C_{16} \leq 5$	Continue current dose.
Improved, but SEs are intolerable	Continue current dose and address SEs, *or* Decrease dose and continue for 2 additional weeks, *or* Go to the next level.
Not improved and SEs are intolerable	Go to the next level.
Return to clinic:	Return in 3 weeks.

FIGURE 21–8. Critical decision point 3 (level 2).

CDP, Week 9	STAR*D Level 2: Citalopram + Bupropion XL
Symptom Improvement (SEs tolerable)	
QIDS-C_{16} ≥ 9	Go to the next level
QIDS-C_{16} = 6–8	Increase dose to 450 mg/day if not previously done, *or* Go to the next level
QIDS-C_{16} ≤ 5	Continue current dose
SEs are intolerable	Go to the next level
Return to clinic:	Return in 3 weeks

FIGURE 21–9. Critical decision point 4 (level 2).

ciated symptoms. Given that the patient's QIDS-C_{16} = 2 and the FIBSER remained acceptable, she was continued on prior medication dosages (Figure 21–9) and rescheduled for follow-up CDP visit in 3 weeks.

CDP 5 (Week 12)

The patient returned for her week 12 visit on time (Figure 21–10). The QIDS-C_{16} was scored as a 2, the FIBSER revealed no side effects, and the patient reported 100% medication adherence. Continued remission into week 12 represents the end of acute phase treatment and signifies moving the patient into the continuation phase of treatment.

Continuation Phase

The continuation phase is the next stage in the pharmacological treatment of MDD. Classification in the continuation phase dictates that the patient achieve full remission of depressive symptoms. The current recommendation for the duration in which psychopharmacological medications are continued in this phase of treatment is a minimum of 6 to 9 months (Keller 2002). In addition to pharmacological treatments, increasing evidence supports the use of evidence-based psychotherapies in the continuation phase of treatment to reduce the risk of relapse (Hollon et al. 2005; Jarrett et al. 2001). Additionally, Fava and colleagues (2004) showed that cognitive-behavioral therapy that sequentially follows a successful course of pharmacotherapy reduces the rates of MDD relapse. With specific regard to this case, after the patient completed participation in the research study, she was referred back to her primary care physician for continued medication management, and follow-up with a cognitive therapist was arranged to help reduce the risk of relapse.

CDP, Week 12	STAR*D Level 2: Citalopram + Bupropion XL
Symptom Improvement (SEs tolerable)	
QIDS-C_{16} ≥ 9	Go to the next level
QIDS-C_{16} = 6–8	Continue current dose and go to follow-up,[1] *or* Go to the next level
QIDS-C_{16} ≤ 5	Continue current dose and go to follow-up
SEs are intolerable	Go to the next level
Return to clinic:	If remitted, return in 3 months or as needed If starting new level, return in 2 weeks

FIGURE 21–10. Critical decision point 5 (level 2).

[1]Because of the severity level of some patients, clinicians may choose to keep patient on current dose and move to follow-up if at week 12 the patient has maintained a score of 6–8 on the QIDS-C_{16} or if the patient does not want to change medications and/or the patient is satisfied with the level of improvement.

Personalized Treatment and Disease Self-Management

Another important component of clinical care with significant promise in the continuing treatment of MDD is counseling patients on disease self-management (Trivedi et al. 2007a). In the case described above, the patient was given a copy of the self-report QIDS-SR_{16} (www.ids-qids.org) with instructions to routinely check for return or worsening of depressive symptoms and contact information for follow-up. Providing patients with the components of MBC are in line with the National Institute of Mental Health's (NIMH) new strategic initiative to target and personalize the treatment of individuals with mental illness. In the plan, Strategy 3.2 aims to "expand and deepen the focus to personalize intervention research." As a part of this strategy, it is suggested that traditional outcome measures used in clinical efficacy studies be expanded to include functional measures and "other indicators of recovery," which must include targeted symptoms and specific symptom profiles. It is hoped that this more personalized treatment will provide patients with depression and other mental illnesses with more thorough recovery.

In conclusion, MBC converges with the NIMH's initiative to personalize treatment and provides patients with education regarding their depressive illness, thereby promoting disease self-management.

Summary

Results from STAR*D indicate the majority of patients will not achieve full remission of depressive symptoms by the end of their first treatment (Trivedi et al. 2006b). On the basis of our clinical experience, determining what subsequent treatments to initiate should be done systematically while paying close attention to objective measures of symptom severity and side-effect burden. These domains constitute the treatment construct known as measurement-based care (Trivedi and Daly 2007; Trivedi et al. 2006b, 2007b). In research settings, measuring these domains is accomplished with the aid of a research assistant; however, in routine clinical practice the same can be done with a nurse or an administrative staff member. The FIBSER is a self-rated measure of side effects, and the QIDS has both a clinician version and a self-rated version that the patient can fill out in the waiting room, with the results easily conveyed to the clinician. The primary goal of MBC is to individualize antidepressant treatment and dosing in an effort to minimize side-effect burden and maintain safety, while enhancing the therapeutic efficacy for each patient. Utilizing a treatment strategy such as MBC in clinical practice provides objective evidence to address key antidepressant treatment questions, such as: At what dose, for what duration, and when do I switch or augment? In addition, MBC can provide ongoing monitoring of symptomatic improvement and side-effect burden in the continuation and maintenance phases of treatment.

References

American Psychiatric Association: Practice guideline for the treatment of patients with major depressive disorder (revision). Am J Psychiatry 157 (suppl 4):1–45, 2000. Available at: http://www.psychiatryonline.com/pracGuide/pracGuideTopic_7.aspx. Accessed June 20, 2009

Biggs MM, Shores-Wilson K, Rush AJ, et al: A comparison of alternative assessments of depressive symptom severity: a pilot study. Psychiatry Res 96:269–279, 2000

Crismon ML, Trivedi M, Pigott TA, et al: The Texas Medication Algorithm Project: report of the Texas Consensus Conference Panel on Medication Treatment of Major Depressive Disorder. J Clin Psychiatry 60:142–156, 1999

Fava GA, Ruini C, Rafanelli C, et al: Six-year outcome of cognitive behavior therapy for prevention of recurrent depression. Am J Psychiatry 161:1872–1876, 2004

Hansen RA, Gartlehner G, Lohr KN, et al: Efficacy and safety of second-generation antidepressants in the treatment of major depressive disorder. Ann Intern Med 143:415–426, 2005

Hollon SD, DeRubeis RJ, Shelton RC, et al: Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. Arch Gen Psychiatry 62:417–422, 2005

Jarrett RB, Kraft D, Doyle J, et al: Preventing recurrent depression using cognitive therapy with and without a continuation phase: a randomized clinical trial. Arch Gen Psychiatry 58:381–388, 2001

Keller MB: Long-term treatment strategies in affective disorders. Psychopharmacol Bull 36 (suppl 2):36–48, 2002

Kroenke K, Spitzer RL, Williams JB: The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 16:606–613, 2001

Paykel ES, Ramana R, Cooper Z, et al.: Residual symptoms after partial remission: an important outcome in depression. Psychol Med 25:1171–1180, 1995

Rush AJ, Trivedi MH, Ibrahim HM, et al: The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 54:573–583, 2003

Rush AJ, Fava M, Wisniewski SR, et al: Sequenced Treatment Alternatives to Relieve Depression (STAR*D): rationale and design. Control Clin Trials 25:119–142, 2004

Rush AJ, Carmody TJ, Ibrahim HM, et al: Comparison of self-report and clinician ratings on two inventories of depressive symptomatology. Psychiatr Serv 57:829–837, 2006a

Rush AJ, Kraemer HC, Sackeim HA, et al: Report by the ACNP Task Force on response and remission in major depressive disorder. Neuropsychopharmacology 31:1841–1853, 2006b

Steer RA, Ball R, Ranieri WF, et al: Dimensions of the Beck Depression Inventory–II in clinically depressed outpatients. J Clin Psychol 55:117–128, 1999

Trivedi MH, Daly EJ: Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. Drug Alcohol Depend 88 (suppl 2):S61–S71, 2007

Trivedi MH, Rush AJ, Ibrahim HM, et al: The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol Med 34:73–82, 2004

Trivedi MH, Fava M, Wisniewski SR, et al: Medication augmentation after the failure of SSRIs for depression. N Engl J Med 354:1243–1252, 2006a

Trivedi MH, Rush AJ, Wisniewski SR, et al: Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry 163:28–40, 2006b

Trivedi MH, Lin EH, Katon WJ: Consensus recommendations for improving adherence, self-management, and outcomes in patients with depression. CNS Spectr 12:1–27, 2007a

Trivedi MH, Rush AJ, Gaynes BN, et al: Maximizing the adequacy of medication treatment in controlled trials and clinical practice: STAR(*)D measurement-based care. Neuropsychopharmacology 32:2479–2489, 2007b

U.S. Food and Drug Administration, Center for Drug Evaluation and Research: Antidepressant use in children, adolescents and adults. March 22, 2004. Available at: http://www.fda.gov/cder/drug/antidepressants/default.htm. Accessed May 2, 2007.

Wisniewski SR, Rush AJ, Balasubramani GK, et al: Self-rated global measure of the frequency, intensity, and burden of side effects. J Psychiatr Pract 12:71–79, 2006

This page intentionally left blank

22

Major Depressive Disorder, Severe With Psychotic Features

Anthony J. Rothschild, M.D.

Setting

I practice in a psychiatry outpatient clinic at a university medical center. The majority of the patients in my practice have major depression. Although like most psychiatrists my role in the care of patients is the prescription of pharmacotherapy, I am a trained psychotherapist, and I do formal psychotherapy with some patients. Even for patients whom I see for "psychopharmacology visits," I provide some psychotherapy during our sessions (i.e., I am not just handing the patients an envelope with information and instructions as to how to take their pills).

Illustration

This study illustrates:

- Difficulty diagnosing psychotic symptoms in major depression
- Decision process involved for electroconvulsive therapy (ECT) versus medications
- Issues in continuation and maintenance therapy: risk of relapse versus medication side effects

Chief Complaint

Mr. P.W. is a 49-year-old male engineer who presents with a chief complaint of "I am very depressed." He was referred by his primary care doctor.

Present Illness

Mr. P.W. stated that about 4–5 months ago he began to feel depressed and "lost interest in life." He thought that the onset of his depression had coincided with a new supervisor at work with whom he had been having trouble getting along. He reported that he feels depressed every day for most of the day. He has lost 9 lb over the past 2 months (current weight, 161 lb), without dieting, because he did not feel like eating. He reported that he has trouble falling asleep almost every night and wakes up at 3:30 A.M. 3–4 nights a week (he normally gets up at 6:30 A.M.). He said he has had no energy for the past 2 months and that it is a struggle to get himself out of bed and to work each day. Once at work, he has trouble completing tasks and concentrating and will frequently sit at his desk and stare out the window for hours at a time. This, he believes, is part of the reason he recently received a poor quarterly performance review at work. He denied suicidal ideation presently but acknowledged that he occasionally feels like "the world would be better off if he died in his sleep." He denied ever having thought of ways to harm himself.

Other Significant Findings From Assessment

Mr. P.W. denied any previous psychiatric history. He has one brother with depression who he thinks responded to sertraline, but he does not know at what dosage. He denied any history of hypomania or

mania. He has a glass of wine with dinner and has another glass of wine before he goes to bed. He has been married for 20 years, and he and his wife have two children, ages 18 and 16. He has been employed for 10 years with his current company, where he is involved with the design and production of weapons systems for the United States military. He denied drug use. He said he does not exercise. Physical examinations and the results of laboratory testing including complete blood count, electrolytes, blood urea nitrogen, creatinine, calcium, glucose, thyroid function tests, folate, and vitamin B_{12} were within normal limits. Mr. P.W.'s primary care doctor had ordered a head magnetic resonance imaging scan because Mr. P.W. had thought he had a brain tumor and insisted that a scan be done. The scan was normal. On mental status examination, Mr. P.W. was cooperative and psychomotorically retarded. He answered most questions with short answers, often simply saying "yes" or "no." He denied symptoms of paranoia or hallucinations. He reported no unusual thoughts or delusions. He was worried that his inability to concentrate and the poor performance review he received might result in his losing his job. He reported that he is still going to work.

Differential Diagnosis

Mr. P.W. meets DSM-IV-TR (American Psychiatric Association 2000a) criteria for major depressive disorder. He exhibits five (or more) symptoms of major depression (depressed mood, loss of interest or pleasure, insomnia, weight loss, trouble concentrating, loss of energy, psychomotor retardation, and thoughts of death), which have been present during the same 2-week period and represent a change from previous functioning. The symptoms are not due to a general medical condition. There is nothing in Mr. P.W.'s history to suggest he has bipolar disorder.

DSM-IV-TR Diagnosis

Axis I	Major depressive disorder, single episode, severe, without psychotic features (296.23)
Axis II	None
Axis III	None
Axis IV	Work stress
Axis V	GAF score: 55

Treatment Plan Considerations

Major Depression, Single Episode

On first assessment Mr. P.W. appears to have a fairly straightforward case of major depressive disorder. He has many symptoms, which indicates that this is a severe episode. His depressive symptoms appear to have affected his ability to function particularly at work. However, he is still going to work. He does not have suicidal ideation at initial assessment and has no active plans for suicide.

Because he is not actively suicidal, homicidal, or psychotic and is still going to work, hospitalization is not indicated, and I decide to treat him as an outpatient. According to the American Psychiatric Association's current practice guideline for the treatment of patients with major depressive disorder (American Psychiatric Association 2000b), in the acute phase the psychiatrist may choose between several initial treatment modalities, including pharmacotherapy, psychotherapy, the combination of medications plus psychotherapy, or ECT. Given that this is Mr. P.W.'s first episode of depression and that he is not actively suicidal or psychotic, I do not believe that ECT is indicated. According to the guideline, because he meets criteria for a severe episode of major depressive disorder Mr. P.W. should be prescribed medication unless he is receiving ECT. I discuss with him the options of medication, psychotherapy, or the combination, and he declines psychotherapy because he feels he does not have enough time or interest to pursue it. He would like to be treated with an antidepressant.

Normally I would review with the patient the previous antidepressant trials received, whether the patient had responded, and whether he or she experienced any side effects. In the case of Mr. P.W., he has never been treated with an antidepressant, so in theory one could prescribe any number of possible antidepressants. However, because he reported that his brother had responded to sertraline, I suggest that he should try sertraline as well. Although there is no firm evidence that if a first-degree relative has responded to a particular antidepressant, it is more likely that the patient will respond as well, in my experience it does seem to increase the likelihood of a response, perhaps just for psychological reasons. Therefore, after reviewing the possible side effects of sertraline with Mr. P.W. and discussing with him how to take the medication, I prescribe sertraline

50 mg/day to be taken in the morning. I advise him to give me a phone call in 3–4 days to report how he is tolerating the medication and have him schedule an appointment to see me again in 10 days.

Treatment Goals, Measures, and Methods

The initial goal for Mr. P.W. is to be able to tolerate the antidepressant medication so he can stay on it long enough to have an antidepressant response. That is the first goal, and it is an important one given that noncompliance with antidepressant medication is a big problem (Melfi et al. 1998). That is one of the reasons I ask patients to call me a few days after starting the medication. Another reason is to make sure that the patient's illness is not getting worse. I will also be monitoring Mr. P.W.'s depressive symptoms with the Hamilton Rating Scale for Depression (Ham-D; Hamilton 1960). In some situations I may formally administer the scale; in others, I ask the same questions that are on the scale during my clinical interview. For Mr. P.W., I will use the Ham-D.

Course

Phone Call (Day 4)

Mr. P.W. calls to report that he has had no side effects from the medication. He also reports that he called in sick to work today because he did not have the energy to get out of bed but plans to go to work tomorrow.

Visit 2 (10 Days)

Mr. P.W. returns for his second visit accompanied by his wife. She asks if she can speak to me. With Mr. P.W.'s permission, she joins the beginning of the appointment. Mr. P.W. reports that he is unchanged and still feels depressed. He again denies symptoms of hallucinations, suicidal ideation, or unusual thoughts, and no delusions are noted. At this point, his wife urges him to "Tell Dr. Rothschild about the car across the street," whereupon the patient described a parked car that he believed contained FBI agents on a stakeout who were following him everywhere he went. The patient then explained a rather detailed delusional system that he was being targeted as part of an investigation of a foreign country's attempt to purchase military secrets. Although this was seemingly plausible given his line of work, his wife reported that neither her husband nor her husband's company worked on classified information requiring a security clearance and, in any case, the cars he was worried about belonged to neighbors. Mr. P.W. had answered in the negative to questions regarding paranoia and unusual thoughts, because to him, the delusional system was "real" and was not an example of unusual thoughts or a reflection of paranoia.

Revised DSM-IV-TR Assessment

After the second visit with the patient and his wife, it became clear that Mr. P.W. had major depressive disorder, single episode, severe with psychotic features (296.24). When I presented my reassessment to Mr. P.W., he was adamant that I was mistaken. In part, this was because the word "psychotic" was frightening to him, and I think it had pejorative connotations for him. Thus I have found using the term "irrational worries" instead of "psychosis" more acceptable to patients and have found its use to result in a greater likelihood of patients telling me what is on their mind (Rothschild 2009). Mr. P.W. felt better with this description of his illness. He, himself, began to refer to his illness as major depressive disorder with "irrational worries."

Axis I	Major depressive disorder, single episode, severe with psychotic features (296.24)
Axis II	None
Axis III	None
Axis IV	Work stress
Axis V	GAF score: 55

Additional Treatment Plan Considerations

Mr. P.W.'s main problem is major depression with psychotic features (psychotic depression), a serious illness during which a person has the dangerous combination of depressed mood and psychosis, with the psychosis commonly manifesting itself as nihilistic delusions that "bad things are about to happen." The American Psychiatric Association's practice guideline for the treatment of patients with major depressive disorder (American Psychiatric Association 2000b) recommends, with substantial clinical confidence, the use of either ECT or the combination of an antipsychotic and an antidepressant for

the treatment of psychotic depression. Several algorithms have recently been proposed, incorporating the current evidence base, to help guide the clinician in the use of somatic treatments for psychotic depression (Hamoda and Osser 2008; Rothschild 2009). Mr. P.W. and his wife were concerned about possible side effects of confusion and memory disruption from ECT and stated that they preferred medication treatment. Because Mr. P.W. was not suicidal, and his current clinical situation was not life threatening, I agreed with their preference for a trial of medication (for further discussion of the decision-making process when deciding between ECT and medications for psychotic depression, see Rothschild 2009).

Of the two recently published algorithms, I followed my own (it is never a good thing not to follow your own published guidelines!), which is based on the evidence base of randomized controlled clinical trials. In the algorithm I point out that there are four combinations of antidepressant plus antipsychotic medications that have been studied and shown to be effective in randomized controlled clinical trials in patients with psychotic depression. These include sertraline plus olanzapine (Meyers et al. 2009; 259 subjects), fluoxetine plus olanzapine (Rothschild et al. 2004; 249 subjects), venlafaxine plus quetiapine (Wijkstra et al. 2008; 122 subjects), and amitriptyline plus perphenazine (Spiker et al. 1985; 51 subjects). Other combinations of antidepressants and antipsychotics have not been studied in randomized controlled clinical trials. After discussion with Mr. P.W. (and, with his permission, his wife), we decided to treat him with a trial of sertraline plus olanzapine. We also decided to dose the medication aggressively, as was done in the National Institute of Mental Health Study of the Pharmacotherapy of Psychotic Depression (STOP-PD; Meyers et al. 2009). The daily dosages of medications in the STOP-PD study (Meyers et al. 2009) were as follows:

1. Initial dosages of 50 mg sertraline/placebo and 5 mg of olanzapine as tolerated
2. Increase the dosage of sertraline/placebo by 50 mg/day and of olanzapine by 5 mg/day every 3 days as tolerated
3. Attempt to achieve minimum dosages of 100 mg/day of sertraline/placebo and 10 mg/day of olanzapine by the end of week 1
4. Increase dosages to 150 mg/day of sertraline/placebo and 15 mg/day of olanzapine during week 2
5. Allow dosages of 200 mg/day of sertraline/placebo and 20 mg/day of olanzapine for residual symptoms beginning in week 3

Because Mr. P.W. is already taking 50 mg/day of sertraline, I advise him to increase the dosage of sertraline to 100 mg/day for 3 days and then to take 150 mg/day. I also start him on olanzapine 5 mg at bedtime (after reviewing possible side effects) and advise him to increase the dosage to 10 mg at bedtime after 3 days. I plan to see him back in 7 days but instruct him to contact me before the appointment if he has trouble tolerating the medications or if his symptoms worsen.

In addition to the pharmacotherapy, I plan to use psychoeducational/cognitive therapy. I use a modified version of cognitive-behavioral therapy based in large part on the work of Gaudiano et al. (2007).

I also advise Mr. P.W. to take a leave of absence from work. This is in part due to the fact that his delusional system involves work. In addition, Mr. P.W. himself admits that he is having trouble functioning and concentrating at work and agrees that a leave of absence is a good idea. He will use accumulated sick time.

Treatment Goals, Measures, and Methods

Table 22–1 outlines my initial treatment goals, measures, and methods.

The initial goal for Mr. P.W. is for him to be able to tolerate the antidepressant and antipsychotic medication combination. In my experience, the psychotic symptoms in psychotically depressed patients can improve quickly, whereas the depressive symptoms may take considerably longer to improve. In the psychotically depressed patient, one can monitor the psychotic symptoms with the Brief Psychiatric Rating Scale (BPRS; Overall and Gorham 1962) and the Delusional Assessment Scale (DAS; Meyers et al. 2006); I will continue to monitor the depressive symptoms with the Ham-D. As mentioned earlier, in some situations I may formally administer the scales, whereas in other situations I ask the same questions that are on the scales during my clinical interview. For Mr. P.W., I will use the Ham-D and the BPRS.

Although the person with an episode of psychotic depression can expect to make a full recovery and re-

TABLE 22–1. Initial treatment goals, measures, and methods

Treatment goal	Measure	Method
Eliminate psychotic symptoms within 10 days	Brief Psychiatric Rating Scale, self-report, probing questions	Pharmacotherapy
Reduce depressive symptoms by 50% or more by 8 weeks	Hamilton Rating Scale for Depression, self-report	Pharmacotherapy Cognitive-behavioral therapy/ psychoeducation
Reduce stress at work	Self-report	Pharmacotherapy (delusional re: work) Use sick time until no longer delusional Return part-time at first Return full-time

turn to full occupational and social functioning in the community, how soon that will happen is highly variable. Some patients will respond quickly to treatment, with fast resolution of depressive and psychotic symptoms, whereas others may have only partial improvement in symptoms, with a significant period of time elapsing until full recovery (Rothschild 2009). In my experience, the patient who is likely to return to full social and occupational functioning quickly is someone who meets the following criteria: first episode, a good response to the somatic treatment prescribed (whether pharmacotherapy or ECT), age between 30 and 65 years, and good premorbid functioning (Rothschild 2009). Mr. P.W. meets several of the criteria for someone likely to have a full recovery quickly.

His Ham-D score is 30, and his BPRS score is 55.

Course (Continued)

Visit 3 (17 Days)

Mr. P.W. reports that he is feeling better. He reports no side effects from the medication. He says he is "no longer concerned about the men in the cars." He states that they are still present across the street, but they are "no longer that interested" in him. I consider this improvement. Although he is still delusional, the "delusions," according to Mr. P.W., are losing interest in him. He reports no change in his depressive symptoms. His Ham-D score is 28, and his BPRS score is 40. I advise him to keep the sertraline at 150 mg/day in the morning and increase the olanzapine to 15 mg at bedtime.

I was torn between holding the olanzapine at 10 mg/day and making the increase to 15 mg/day. An argument for staying at 10 mg/day is that he is improving, whereas arguments for advancing to 15 mg/day include the following:

1. He is still psychotic.
2. The STOP-PD study used 15 mg/day of olanzapine as a target (Meyers et al. 2009).
3. Other studies of psychotic depression that used olanzapine in combination with an antidepressant had a mean dosage of olanzapine between 12.4 and 13.9 mg/day (Rothschild et al. 2004).
4. The serious morbidity and mortality of psychotic depression requires more aggressive treatment.

Visit 4 (24 Days)

Mr. P.W. reports that he has continued to improve. He reports no side effects from the medication except he experienced some morning drowsiness for a few days after the olanzapine was increased to 15 mg/day. He reports no change in appetite or weight. He reports that the "men in the cars" still occasionally drive by his house, but they appear to be doing so less frequently, and he believes they are "losing interest" in him. He still reports a number of depressive symptoms and still feels sad. He is sleeping well. He still has no appetite. His Ham-D score is 16, and his BPRS score is 32. He worries that his poor performance at work and his current leave of absence may cost him his job. He would like to return to work. I advise him that it is still premature to return to work and that it could be counterproductive for him to return to work before he is well. I reassure him that he will get better and that his improvement to date is a good prognostic sign. I advise him to continue his medications at sertraline 150 mg in the morning and olanzapine 15 mg at bedtime.

Visit 5 (31 Days)

For the first time since I have been seeing him, I notice that Mr. P.W. smiles appropriately. He is not aware of this (in my experience, seeing a depressed person smile for the first time when they had not done so before is frequently a harbinger of future improvement). He feels only mildly depressed. No delusional thinking is noted on examination. He says that the people watching him are "gone." His Ham-D score is 11 and his BPRS score is 25. He complains that he has gained 2 lb since he last weighed himself (2 weeks ago) and says that he is eating more. He is concerned that the weight gain is related to his medication. I point out to him that although it is possible that his increased appetite and weight gain are related to the medication, he did lose 9 lb in the 2 months prior to coming to see me and that the increased appetite and weight gain may be related to the improvement in his depression. We will continue to monitor his weight. He again asks about returning to work. I advise that I would prefer to see him have at least another week (2 weeks in total) of no serious depressive symptoms and no "irrational worries." I suggested to him that he contact his supervisor as to whether he could return part-time in about 9–10 days from today (after our next appointment). I also advise continuation of his current medication regimen of sertraline 150 mg/day in the morning and olanzapine 15 mg/day at bedtime.

Phone Call (40 Days)

I receive a phone call from Mr. P.W.'s wife to let me know that he appears to be back to "his old self."

Visit 6 (41 Days)

Improvement has continued. Mr. P.W. denies depression. When I ask about the men in the cars across the street he looks at me with a surprised look and says, "I now know that I had those thoughts because I was ill. I cannot believe I thought that." We then discuss the delusional thoughts, "the irrational worries," he had experienced. It becomes very clear that Mr. P.W. is having a very hard time reconciling that he has had these thoughts. This is a very good sign. When a patient with psychotic depression has insight that the unusual thoughts were part of the illness and almost has a cognitive dissonance that they had once believed them, I am confident that they are in the process of making a full recovery. His Ham-D score is 9, and his BPRS score is 21. Mr. P.W. reports that his supervisor said he could return to work with a note from me. I write a note stating that he can return to work at 50% time for 1 week and full-time after that. (My plan was to see Mr. P.W. again after he had been back at work part-time for 1 week before deciding whether he was ready to return full-time.) I advise continuation of his current medication regimen of sertraline 150 mg/day in the morning and olanzapine 15 mg/day at bedtime.

Visit 7 (Week 9)

Mr. P.W. reports that his week back at work went well and that his supervisor would like him to return to work full-time. He denies any symptoms of depression. His Ham-D score is 7. No delusions are noted on examination. His BPRS score is 20. He complains again about increased appetite and weight gain. Now up to 170 lb, Mr. P.W. has gained 9 lb since I began treating him and is back to his starting weight before he was depressed. I reassure him that we will continue to monitor his weight. I write him a note to return to work full-time. He is to continue on sertraline 150 mg/day in the morning and olanzapine 15 mg/day at bedtime.

Visit 8 (Week 13)

Mr. P.W. is doing well. Work is also going well. He has had no depressive or psychotic symptoms. His Ham-D score is 7, and his BPRS score is 20. His weight is 172 lb.

Visit 9 (Week 17)

He is still doing well. Work is also still going well. No depressive or psychotic symptoms. His Ham-D score is 7, and his BPRS score is 20. His weight is 173 lb.

Visit 10 (Week 21)

He is still doing well. Work is also still going well. No depressive or psychotic symptoms. His Ham-D score is 7, and his BPRS score is 20. His weight is 174 lb.

Visit 11 (Week 25)

He is still doing well. Work is also still going well. No depressive or psychotic symptoms. His Ham-D score is 7, and his BPRS score is 20. His weight is 174 lb.

Visit 12 (Week 29)

He is still doing well. Work is also still going well. No depressive or psychotic symptoms. His Ham-D score is 7, and his BPRS score is 20. His weight is 176 lb. At this point, Mr. P.W. has gained 6 lb above his premorbid weight of 170 lb and has been stable psychiatrically for at least 4 months. Based on studies of the continuation and maintenance treatment of psychotic depression (Rothschild 2009), it has been my practice to leave a patient on the combination of the antidepressant/antipsychotic to which he or she responded to for 4 months. After 4 months, if the patient has continued to remain in remission, I begin a gradual taper of the antipsychotic medication, leaving the patient on the antidepressant. If the patient is having significant side effects (e.g., signs of tardive dyskinesia with an older antipsychotic medication or metabolic syndrome symptoms with a newer antipsychotic agent), I may start the taper earlier than 4 months. On the other hand, if the patient is not having any side effects, and/or is still symptomatic, I may delay the taper of the antipsychotic medication beyond 4 months. I usually will leave the patient on the antidepressant indefinitely, given the high rate of relapse in psychotic depression and the significant morbidity and mortality associated with relapses. Given Mr. P.W.'s current stability for at least 4 months and gradual weight gain, I begin a gradual taper of the olanzapine by 5 mg every 4 weeks. He is to remain on sertraline at 150 mg/day.

Visit 13 (Week 33)

He is still doing well. Work is also still going well. No depressive or psychotic symptoms. He is taking olanzapine 10 mg/day and sertraline 150 mg/day. His Ham-D score is 7, and his BPRS score is 20. His weight is 175 lb.

Visit 14 (Week 37)

He is still doing well. Work is also still going well. No depressive or psychotic symptoms. He is taking olanzapine 5 mg/day and sertraline 150 mg/day. His Ham-D score is 7, and his BPRS score is 20. His weight is 174 lb.

Visit 15 (Week 41)

Mr. P.W. is now taking just sertraline 150 mg/day. He is doing well, and work is going well. No depressive or psychotic symptoms. His Ham-D score is 7, and his BPRS score is 20. His weight is 173 lb.

Visit 16 (Week 46)

Mr. P.W. continues on sertraline 150 mg/day. He is doing well, and work is going well. No depressive or psychotic symptoms. His Ham-D score is 7, and his BPRS score is 20. His weight is 172 lb.

Summary of Guideline Use

I followed the practice guidelines for treatment of the patient with psychotic depression, which recommend either the combination of an antidepressant and an antipsychotic or ECT (American Psychiatric Association 2000b; Hamoda and Osser 2008; Rothschild 2009). A delay in the implementation of these guidelines occurred when the diagnosis of psychotic depression was initially missed because PW's delusional thinking was not readily apparent. This, unfortunately, is a not an infrequent occurrence (Rothschild et al. 2008), even among experts.

Ways to Improve Practice

In reviewing my practice, I could have improved the care of Mr. P.W. by detecting his psychotic features earlier. Had I asked more probing questions or involved his wife earlier, I might have learned earlier about his paranoid delusional thoughts.

I now refer to delusional thinking with my depressed patients as "irrational worries" and find that my patients have an easier time accepting this concept than "psychotic thinking." This then results in their feeling more comfortable about telling me what is worrying them and what is on their mind.

References

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000a

American Psychiatric Association: Practice guideline for the treatment of patients with major depressive disorder, 2nd edition, in Practice Guidelines for the Treatment of Psychiatric Disorders, Compendium. Washington, DC, American Psychiatric Association, 2006, pp 763–840

Gaudiano BA, Miller IW, Hebert JD: The treatment of psychotic major depression: is there a role for adjunctive psychotherapy? Psychother Psychosom 76:271–277, 2007

Hamilton M: A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56–62, 1960

Hamoda HM, Osser DN: The psychopharmacology algorithm project at the Harvard South Shore Program: an update on psychotic depression. Harv Rev Psychiatry 16:235–247, 2008

Melfi CA, Chawla AJ, Croghan TW, et al: The effects of adherence to antidepressant treatment: guidelines on relapse and recurrence of depression. Arch Gen Psychiatry 55:1128–1132, 1998

Meyers B, English J, Gabriele M, et al: A delusion assessment scale for psychotic major depression: reliability, validity and utility. Biol Psychiatry 60:1336–1342, 2006

Meyers BS, Flint AJ, Rothschild AJ, et al: A double-blind randomized controlled trial of olanzapine plus sertraline vs olanzapine plus placebo for psychotic depression: the Study of Pharmacotherapy of Psychotic Depression (STOP-PD). Arch Gen Psychiatry 66:838–847, 2009

Overall JE, Gorham DE: The Brief Psychiatric Rating Scale. Psychol Rep 10:799–812, 1962

Rothschild AJ (ed): Clinical Manual for the Diagnosis and Treatment of Psychotic Depression. Washington, DC, American Psychiatric Publishing, 2009

Rothschild AJ, Williamson DJ, Tohen MF, et al: A double-blind, randomized study of olanzapine and olanzapine/fluoxetine combination for major depression with psychotic features. J Clin Psychopharmacol 24:365–373, 2004

Rothschild AJ, Winer J, Fratoni S, et al: Missed diagnosis of psychotic depression at 4 academic medical centers. J Clin Psychiatry 69:1293–1296, 2008

Spiker DG, Weiss JC, Dealy RS, et al: The pharmacological treatment of delusional depression. Am J Psychiatry 142:430–436, 1985

Wijkstra JJ, Burger, D, van den Broek WW, et al: Pharmacological treatment of psychotic depression: a randomized, double-blind study comparing imipramine, venlafaxine and venlafaxine plus quetiapine. Presented at the annual meeting of the American Psychiatric Association, Washington, DC, May 2008

Patient and Family Resources

Families for Depression Awareness (www.familyaware .org)

Rothschild AJ (ed): Clinical Manual for the Diagnosis and Treatment of Psychotic Depression. Washington, DC, American Psychiatric Publishing, 2009

23

Acute Bipolar Depression

Po W. Wang, M.D.
Terence A. Ketter, M.D.

Setting

The patient was seen in an academic outpatient practice setting. The surrounding population is of high educational attainment and upper socioeconomic status and ethnically diverse. The patient population is sophisticated in their approach to accessing Internet-based information. The treatment strategy used was evidence-based pharmacotherapy with a combination of psychoeducation and cognitive-behavioral therapy (CBT).

Illustration

This study illustrates:

- Importance of differentiating bipolar from unipolar depression
- Use of standardized clinical monitoring forms
- Use of medication tracking forms
- Use of standardized adverse effects monitoring and tracking
- Use of standardized mood severity ratings
- Application of guideline-based treatment for bipolar depression
- Integrating manualized psychoeducation
- Application of manualized CBT

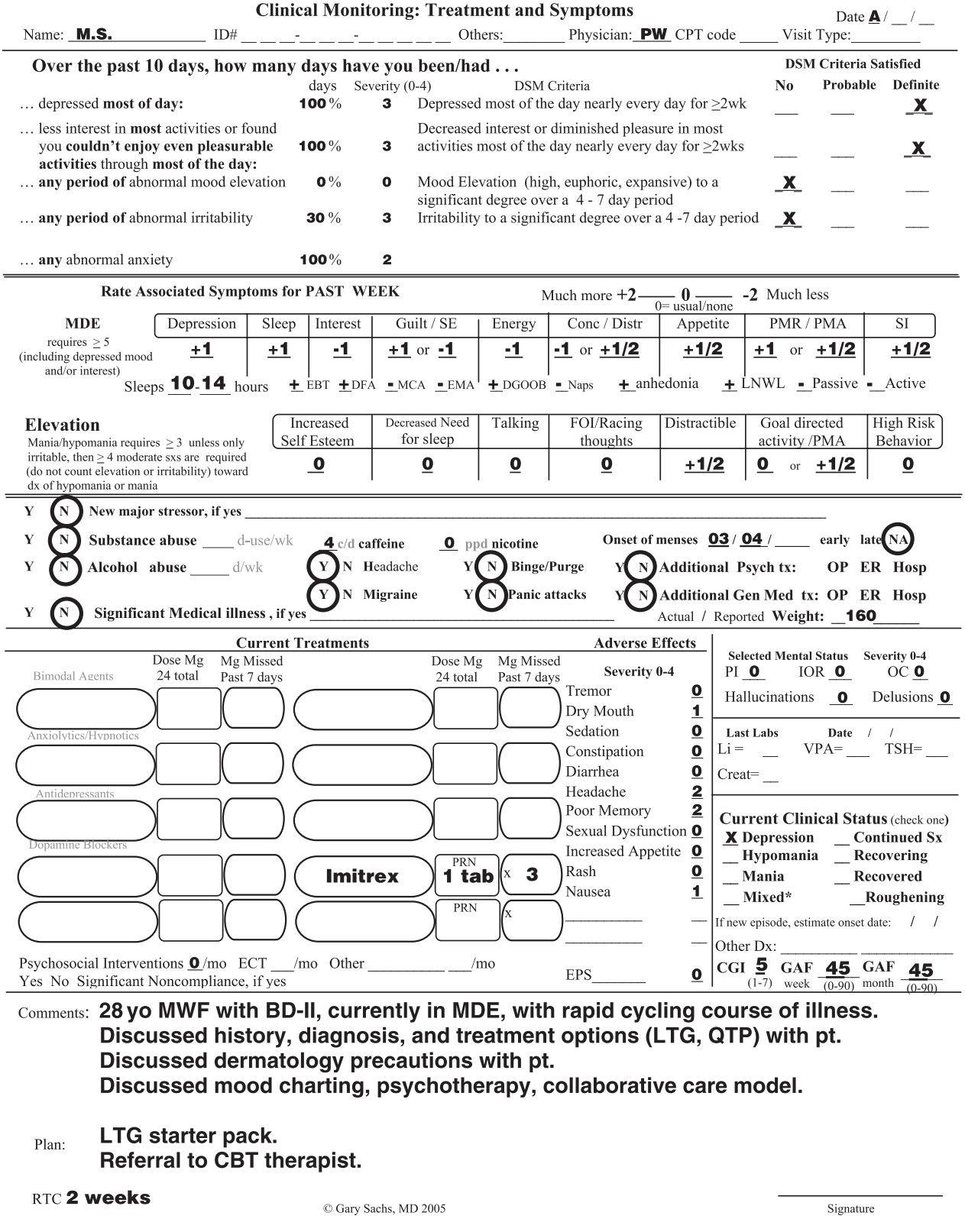
Chief Complaint

Ms. M.S. is a 28-year-old married Caucasian woman who reports a long history of "mood swings" since adolescence and presents with treatment-resistant depression. She arrives for the evaluation with her husband.

Present Illness

The patient reported depressive episodes manifested by depressed mood and anhedonia with social isolation. Despite sleeping 10–14 hours per night, she has low energy and psychomotor retardation during the day. Each of these symptoms meets DSM-IV-TR (American Psychiatric Association 2000) criteria for being pervasively present nearly all day and nearly every day over at least the past 2 weeks prior to her presentation to the clinic. Appetite has increased, but without significant weight gain. She has had thoughts of suicide at times during the week but not any plans or intentions to end her life. She has ruminated on her poor performance in school, past failed relationships, and current financial difficulties. In addition, she has had anxiety with palpitations and mild tremors sporadically, but no shortness of breath, diaphoresis, or nausea. The current episode is 3 weeks in duration. A completed Clinical Monitoring Form is shown in Figure 23–1A.

These episodes began in elementary school, as far as Ms. M.S. could recall. In the past, these episodes ranged in duration from 2 to 4 weeks, followed by 2 to 4 weeks during which her mood returned to a baseline level of poor self-esteem and sleeping 9–10 hours per night, but with good function in school, relationships, and work. Her mood would continuously alternate between these two states of low self-



Clinical Monitoring: Treatment and Symptoms

Date **A** / __ / __

Name: **M.S.** ID# __ __ __-__ __ __-__ __ __ __ __ Others:_____ Physician: **PW** CPT code _____ Visit Type:_____

Over the past 10 days, how many days have you been/had . . .

	days	Severity (0-4)	DSM Criteria	DSM Criteria Satisfied: No	Probable	Definite
… depressed **most of day:**	**100** %	**3**	Depressed most of the day nearly every day for ≥2wk	___	___	**X**
… less interest in **most** activities or found you **couldn't enjoy even pleasurable activities** through **most of the day:**	**100** %	**3**	Decreased interest or diminished pleasure in most activities most of the day nearly every day for ≥2wks	___	___	**X**
… **any period of** abnormal mood elevation	**0** %	**0**	Mood Elevation (high, euphoric, expansive) to a significant degree over a 4 - 7 day period	**X**	___	___
… **any period of** abnormal irritability	**30** %	**3**	Irritability to a significant degree over a 4 -7 day period	**X**	___	___
… **any** abnormal anxiety	**100** %	**2**				

Rate Associated Symptoms for PAST WEEK Much more +2 ____ **0** ____ -2 Much less (0= usual/none)

MDE (requires ≥ 5 (including depressed mood and/or interest))	Depression	Sleep	Interest	Guilt / SE	Energy	Conc / Distr	Appetite	PMR / PMA	SI
	+1	**+1**	**-1**	**+1** or **-1**	**-1**	**-1** or **+1/2**	**+1/2**	**+1** or **+1/2**	**+1/2**

Sleeps **10**-**14** hours **+** EBT **+** DFA **-** MCA **-** EMA **+** DGOOB **-** Naps **+** anhedonia **+** LNWL **-** Passive **-** Active

Elevation
Mania/hypomania requires ≥ 3 unless only irritable, then ≥ 4 moderate sxs are required (do not count elevation or irritability) toward dx of hypomania or mania

Increased Self Esteem	Decreased Need for sleep	Talking	FOI/Racing thoughts	Distractible	Goal directed activity /PMA	High Risk Behavior
0	**0**	**0**	**0**	**+1/2**	**0** or **+1/2**	**0**

Y (N) **New major stressor, if yes** _____

Y (N) **Substance abuse** _____ d-use/wk **4** c/d **caffeine** **0** ppd **nicotine** **Onset of menses 03 / 04 / _____ early late (NA)**

Y (N) **Alcohol abuse** _____ d/wk (Y) N Headache Y (N) **Binge/Purge** Y (N) **Additional Psych tx: OP ER Hosp**

(Y) N Migraine Y (N) **Panic attacks** Y (N) **Additional Gen Med tx: OP ER Hosp**

Y (N) **Significant Medical illness , if yes** _____ Actual / Reported **Weight: __160_____**

Current Treatments

	Dose Mg 24 total	Mg Missed Past 7 days
Bimodal Agents		
Anxiolytics/Hypnotics		
Antidepressants		
Dopamine Blockers		
Imitrex	PRN **1 tab**	x **3**
	PRN	x

Psychosocial Interventions **0**/mo ECT ___/mo Other _____ ___/mo

Yes No Significant Noncompliance, if yes

Adverse Effects

Severity 0-4	
Tremor	**0**
Dry Mouth	**1**
Sedation	**0**
Constipation	**0**
Diarrhea	**0**
Headache	**2**
Poor Memory	**2**
Sexual Dysfunction	**0**
Increased Appetite	**0**
Rash	**0**
Nausea	**1**
_____	—
_____	—
EPS_____	**0**

Selected Mental Status **Severity 0-4**
PI **0** IOR **0** OC **0**
Hallucinations **0** Delusions **0**

Last Labs **Date** / /
Li = __ VPA= ___ TSH= ___
Creat= __

Current Clinical Status (check one)
X Depression __ **Continued Sx**
__ **Hypomania** __ **Recovering**
__ **Mania** __ **Recovered**
__ **Mixed*** __**Roughening**
If new episode, estimate onset date: / /
Other Dx: _____ _____
CGI 5 (1-7) **GAF 45** week (0-90) **GAF 45** month (0-90)

Comments: **28 yo MWF with BD-II, currently in MDE, with rapid cycling course of illness.**
Discussed history, diagnosis, and treatment options (LTG, QTP) with pt.
Discussed dermatology precautions with pt.
Discussed mood charting, psychotherapy, collaborative care model.

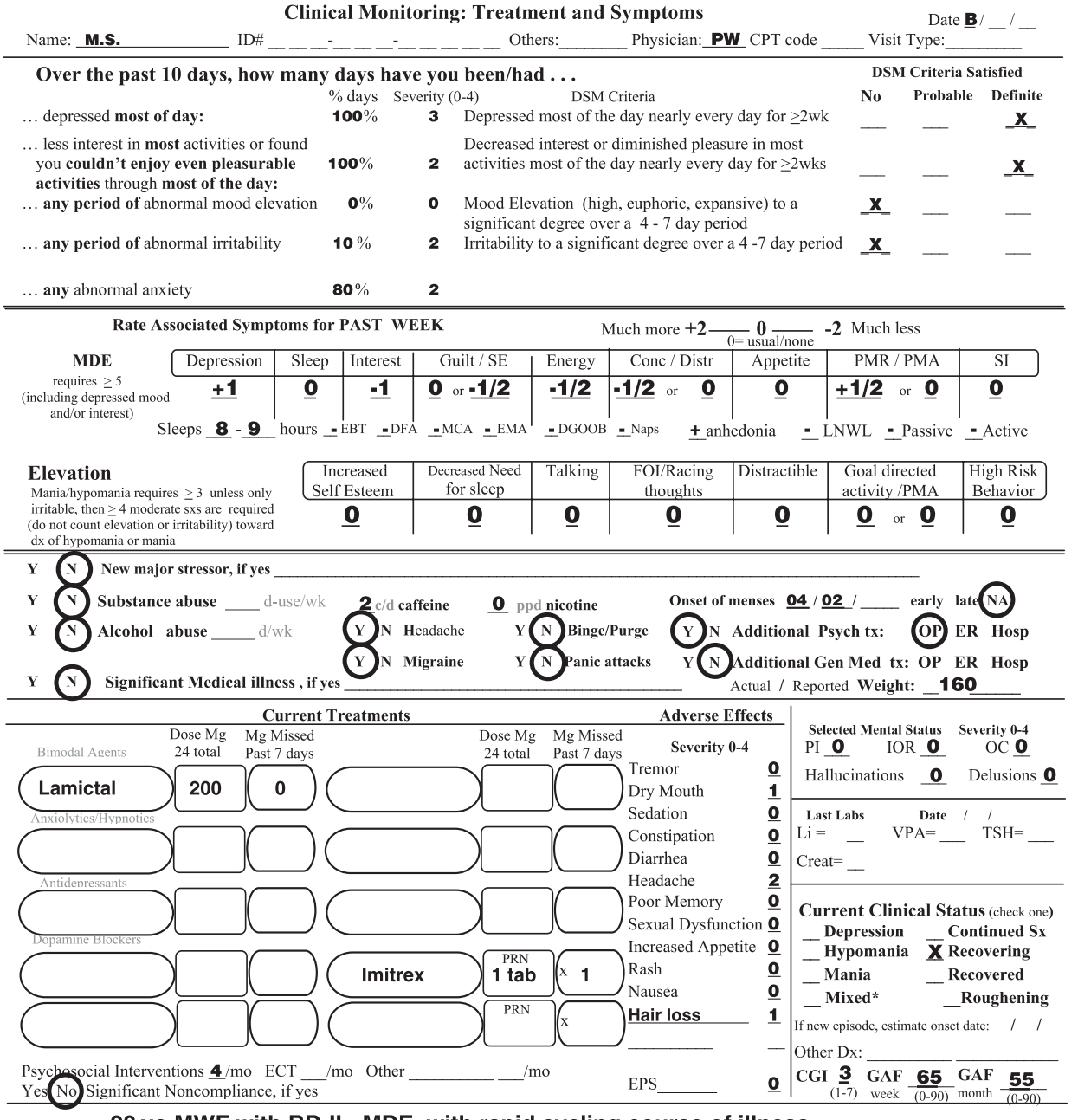
Plan: **LTG starter pack.**
Referral to CBT therapist.

RTC **2 weeks**

Signature

FIGURE 23–1. Clinical monitoring form–representative time points (A).

A, Initial intake. B, Four weeks after starting lamotrigine. C, Regular follow-up visit during 2½ years of euthymic mood. D, Subsyndromal mood elevation during euthymic mood period.



Clinical Monitoring: Treatment and Symptoms

Date **B** / __ / __

Name: **M.S.** ID# __ __ __-__ __ __-__ __ __ __ __ Others:_____ Physician: **PW** CPT code _____ Visit Type:_____

Over the past 10 days, how many days have you been/had . . .

	% days	Severity (0-4)	DSM Criteria	DSM Criteria Satisfied: No	Probable	Definite
… depressed **most of day:**	**100**%	**3**	Depressed most of the day nearly every day for ≥2wk	___	___	**X**
… less interest in **most** activities or found you **couldn't enjoy even pleasurable activities** through **most of the day:**	**100**%	**2**	Decreased interest or diminished pleasure in most activities most of the day nearly every day for ≥2wks	___	___	**X**
… **any period of** abnormal mood elevation	**0**%	**0**	Mood Elevation (high, euphoric, expansive) to a significant degree over a 4 - 7 day period	**X**	___	___
… **any period of** abnormal irritability	**10** %	**2**	Irritability to a significant degree over a 4 -7 day period	**X**	___	___
… **any** abnormal anxiety	**80**%	**2**				

Rate Associated Symptoms for PAST WEEK Much more **+2** ____ **0** ____ **-2** Much less (0= usual/none)

MDE requires ≥ 5 (including depressed mood and/or interest)

Depression	Sleep	Interest	Guilt / SE	Energy	Conc / Distr	Appetite	PMR / PMA	SI
+1	**0**	**-1**	**0** or **-1/2**	**-1/2**	**-1/2** or **0**	**0**	**+1/2** or **0**	**0**

Sleeps **8** - **9** hours **-** EBT **-** DFA **-** MCA **-** EMA **-** DGOOB **-** Naps **+** anhedonia **-** LNWL **-** Passive **-** Active

Elevation Mania/hypomania requires ≥ 3 unless only irritable, then ≥ 4 moderate sxs are required (do not count elevation or irritability) toward dx of hypomania or mania

Increased Self Esteem	Decreased Need for sleep	Talking	FOI/Racing thoughts	Distractible	Goal directed activity /PMA	High Risk Behavior
0	**0**	**0**	**0**	**0**	**0** or **0**	**0**

Y (N) **New major stressor, if yes** _____

Y (N) **Substance abuse** ____ d-use/wk **2** c/d **caffeine** **0** ppd **nicotine** **Onset of menses 04 / 02 / ____ early late (NA)**

Y (N) **Alcohol abuse** _____ d/wk (Y) N **Headache** Y (N) **Binge/Purge** Y (N) **Additional Psych tx:** (OP) **ER Hosp**

(Y) N **Migraine** Y (N) **Panic attacks** Y (N) **Additional Gen Med tx: OP ER Hosp**

Y (N) **Significant Medical illness , if yes** _____ Actual / Reported **Weight: 160**

Current Treatments

Bimodal Agents	Dose Mg 24 total	Mg Missed Past 7 days		Dose Mg 24 total	Mg Missed Past 7 days
Lamictal	200	0			
Anxiolytics/Hypnotics					
Antidepressants					
Dopamine Blockers					
			Imitrex	PRN **1 tab**	x 1
				PRN	x

Psychosocial Interventions **4** /mo ECT ___/mo Other _____ ___/mo

Yes (No) Significant Noncompliance, if yes

Adverse Effects Severity 0-4

Tremor	**0**
Dry Mouth	**1**
Sedation	**0**
Constipation	**0**
Diarrhea	**0**
Headache	**2**
Poor Memory	**0**
Sexual Dysfunction	**0**
Increased Appetite	**0**
Rash	**0**
Nausea	**0**
Hair loss	**1**
_____	—
EPS_____	**0**

Selected Mental Status **Severity 0-4**

PI **0** IOR **0** OC **0**

Hallucinations **0** Delusions **0**

Last Labs **Date** / /

Li = __ VPA= ___ TSH= ___

Creat= __

Current Clinical Status (check one)

__ **Depression** __ **Continued Sx**

__ **Hypomania** **X** **Recovering**

__ **Mania** __ **Recovered**

__ **Mixed*** __**Roughening**

If new episode, estimate onset date: / /

Other Dx: _____ _____

CGI 3 (1-7) **GAF 65** week (0-90) **GAF 55** month (0-90)

Comments: **28 yo MWF with BD-II - MDE, with rapid cycling course of illness. Improvement noted by family and therapist; pt endorses on prompting. Tolerated LTG titration.**

Started weekly CBT therapy.

Plan:

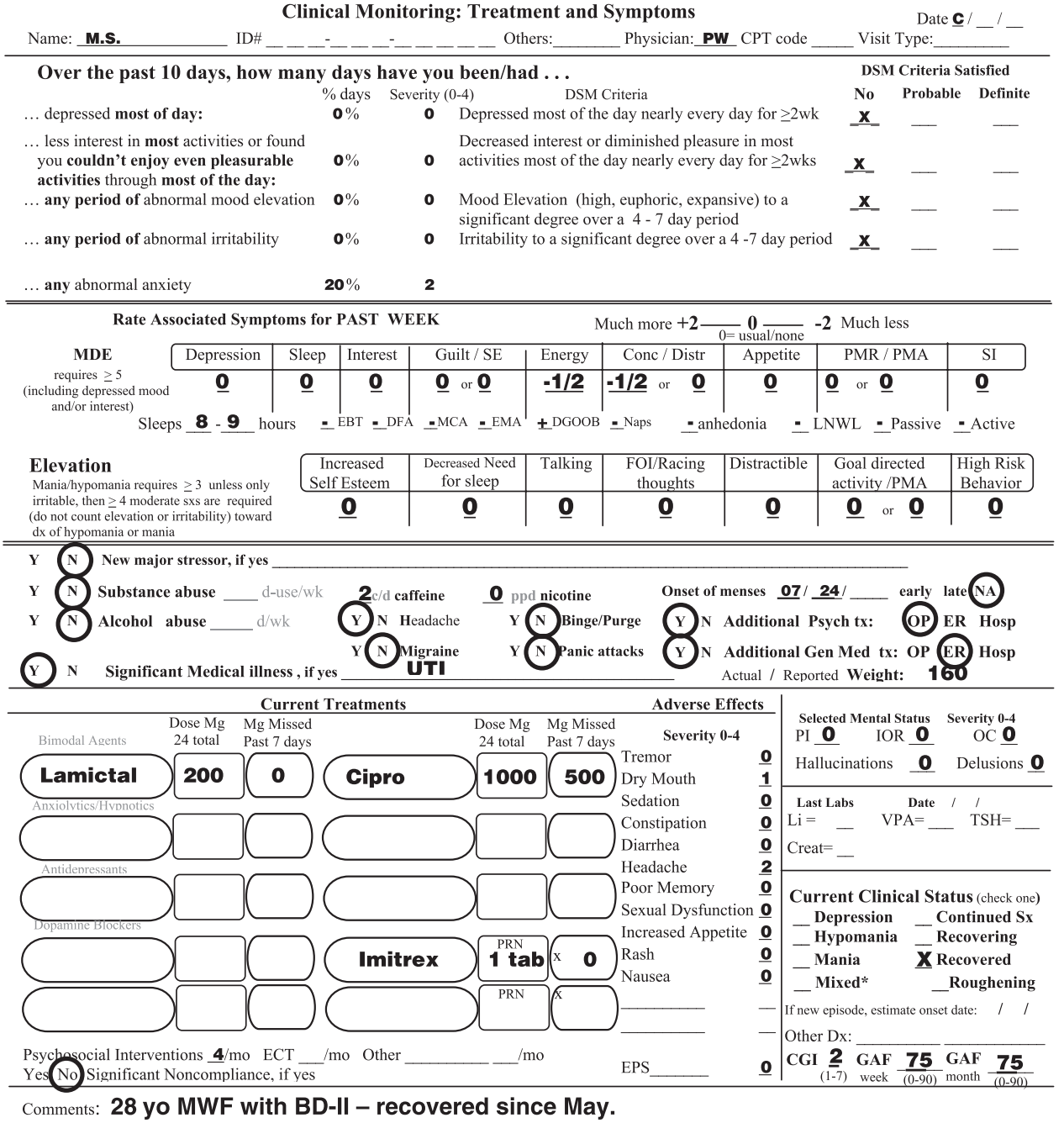
Meds status quo.

RTC **4 weeks**

Signature

FIGURE 23–1. Clinical monitoring form–representative time points (B).

A, Initial intake. **B,** Four weeks after starting lamotrigine. C, Regular follow-up visit during 2½ years of euthymic mood. D, Subsyndromal mood elevation during euthymic mood period.



Clinical Monitoring: Treatment and Symptoms

Date C / __ / __

Name: M.S. ID# __ __ __-__ __ __-__ __ __ __ __ Others:_____ Physician: PW CPT code _____ Visit Type:_____

Over the past 10 days, how many days have you been/had . . .

	% days	Severity (0-4)	DSM Criteria	DSM Criteria Satisfied: No	Probable	Definite
… depressed **most of day:**	0%	0	Depressed most of the day nearly every day for ≥2wk	X		
… less interest in **most** activities or found you **couldn't enjoy even pleasurable activities** through **most of the day:**	0%	0	Decreased interest or diminished pleasure in most activities most of the day nearly every day for ≥2wks	X		
… **any period of** abnormal mood elevation	0%	0	Mood Elevation (high, euphoric, expansive) to a significant degree over a 4 - 7 day period	X		
… **any period of** abnormal irritability	0%	0	Irritability to a significant degree over a 4 -7 day period	X		
… **any** abnormal anxiety	20%	2				

Rate Associated Symptoms for PAST WEEK Much more +2 —— 0 —— -2 Much less (0= usual/none)

MDE (requires ≥ 5 (including depressed mood and/or interest))	Depression	Sleep	Interest	Guilt / SE	Energy	Conc / Distr	Appetite	PMR / PMA	SI
	0	0	0	0 or 0	-1/2	-1/2 or 0	0	0 or 0	0

Sleeps 8 - 9 hours - EBT - DFA - MCA - EMA + DGOOB - Naps - anhedonia - LNWL - Passive - Active

Elevation Mania/hypomania requires ≥ 3 unless only irritable, then ≥ 4 moderate sxs are required (do not count elevation or irritability) toward dx of hypomania or mania

Increased Self Esteem	Decreased Need for sleep	Talking	FOI/Racing thoughts	Distractible	Goal directed activity /PMA	High Risk Behavior
0	0	0	0	0	0 or 0	0

Y (N) **New major stressor, if yes** _____

Y (N) **Substance abuse** ____ d-use/wk — 2 c/d **caffeine** — 0 ppd **nicotine** — **Onset of menses** 07 / 24 / ____ early late (NA)

Y (N) **Alcohol abuse** _____ d/wk — Y (N) **Headache** — Y (N) **Binge/Purge** — Y (N) **Additional Psych tx:** (OP) ER Hosp

Y (N) **Migraine** — Y (N) **Panic attacks** — Y (N) **Additional Gen Med tx:** OP (ER) Hosp

(Y) N **Significant Medical illness , if yes** UTI — Actual / Reported **Weight:** 160

Current Treatments

	Dose Mg 24 total	Mg Missed Past 7 days
Bimodal Agents: **Lamictal**	200	0
Anxiolytics/Hypnotics:		
Antidepressants:		
Dopamine Blockers:		

	Dose Mg 24 total	Mg Missed Past 7 days
Cipro	1000	500
Imitrex	PRN 1 tab	x 0
	PRN	x

Psychosocial Interventions 4/mo ECT ___/mo Other _____ ___/mo

Yes (No) Significant Noncompliance, if yes

Adverse Effects

Severity 0-4	
Tremor	0
Dry Mouth	1
Sedation	0
Constipation	0
Diarrhea	0
Headache	2
Poor Memory	0
Sexual Dysfunction	0
Increased Appetite	0
Rash	0
Nausea	0
_____	—
_____	—
EPS_____	0

Selected Mental Status — **Severity 0-4**: PI 0 IOR 0 OC 0; Hallucinations 0 Delusions 0

Last Labs — **Date** / /: Li = __ VPA= ___ TSH= ___ Creat= __

Current Clinical Status (check one)

__ Depression __ Continued Sx
__ Hypomania __ Recovering
__ Mania X Recovered
__ Mixed* __Roughening

If new episode, estimate onset date: / /

Other Dx: _____ _____

CGI 2 (1-7) GAF week 75 (0-90) GAF month 75 (0-90)

Comments: **28 yo MWF with BD-II – recovered since May.**

Plan: **Meds status quo. Use CBT for residual anxiety.**

RTC **2-3 months**

Signature

FIGURE 23–1. Clinical monitoring form–representative time points (C).

A, Initial intake. B, Four weeks after starting lamotrigine. **C,** Regular follow-up visit during 2½ years of euthymic mood. D, Subsyndromal mood elevation during euthymic mood period.

Clinical Monitoring: Treatment and Symptoms

Date **D** / **09**/ **12**

Name: **M.S.** ID# __ __ __-__ __ __-__ __ __ __ __ Others:_____ Physician: **PW** CPT code ____ Visit Type:_____

Over the past 10 days, how many days have you been/had . . .

	% days	Severity (0-4)	DSM Criteria	DSM Criteria Satisfied: No	Probable	Definite
… depressed **most of day:**	**30**%	**1**	Depressed most of the day nearly every day for ≥2wk	**X**		
… less interest in **most** activities or found you **couldn't enjoy even pleasurable activities** through **most of the day:**	**0**%	**0**	Decreased interest or diminished pleasure in most activities most of the day nearly every day for ≥2wks	**X**		
… **any period of** abnormal mood elevation	**0**%	**0**	Mood Elevation (high, euphoric, expansive) to a significant degree over a 4 - 7 day period	**X**		
… **any period of** abnormal irritability	**40**%	**2**	Irritability to a significant degree over a 4 -7 day period			**X**
… **any** abnormal anxiety	**50**%	**3**				

Rate Associated Symptoms for PAST WEEK Much more +2 —— **0** —— -2 Much less (0= usual/none)

MDE (requires ≥ 5 (including depressed mood and/or interest))	Depression	Sleep	Interest	Guilt / SE	Energy	Conc / Distr	Appetite	PMR / PMA	SI
	+1/2	**-1**	**0**	**0** or **0**	**0**	**0** or **+1/2**	**0**	**0** or **+1**	**0**

Sleeps **6** - **8** hours ▪ EBT **+** DFA ▪ MCA ▪ EMA ▪ DGOOB ▪ Naps ▪ anhedonia ▪ LNWL ▪ Passive ▪ Active

Elevation Mania/hypomania requires ≥ 3 unless only irritable, then ≥ 4 moderate sxs are required (do not count elevation or irritability) toward dx of hypomania or mania

Increased Self Esteem	Decreased Need for sleep	Talking	FOI/Racing thoughts	Distractible	Goal directed activity /PMA	High Risk Behavior
0	**-1**	**+1/2**	**+1/2**	**+1/2**	**0** or **+1**	**0**

Y (N) **New major stressor, if yes** _____

Y (N) **Substance abuse** ____ d-use/wk **2** c/d **caffeine** **0** ppd **nicotine** **Onset of menses** **09**/ **1**/ ____ early late (NA)

Y (N) **Alcohol abuse** _____ d/wk Y (N) **Headache** Y (N) **Binge/Purge** (Y) N **Additional Psych tx:** (OP) ER Hosp

Y (N) **Migraine** Y (N) **Panic attacks** Y (N) **Additional Gen Med tx:** OP ER Hosp

Y (N) **Significant Medical illness , if yes** _____ Actual / Reported **Weight:** **155**

Current Treatments

	Dose Mg 24 total	Mg Missed Past 7 days		Dose Mg 24 total	Mg Missed Past 7 days
Bimodal Agents					
Lamictal	**200**	**400**	**OCP**	**1 tab**	**1 tab**
Anxiolytics/Hypnotics					
Antidepressants					
Dopamine Blockers					
			Imitrex	PRN **1 tab**	x **0**
			Ambien	PRN **10**	x **2**

Psychosocial Interventions **2**/mo ECT ___/mo Other _____ ___/mo

Yes (No) Significant Noncompliance, if yes

Adverse Effects (Severity 0-4)

Tremor	**2**
Dry Mouth	**2**
Sedation	**0**
Constipation	**0**
Diarrhea	**0**
Headache	**0**
Poor Memory	**0**
Sexual Dysfunction	**0**
Increased Appetite	**0**
Rash	**0**
Nausea	**0**
_____	—
_____	—
EPS_____	**0**

Selected Mental Status PI **0** IOR **0** **Severity 0-4** OC **0**

Hallucinations **0** Delusions **0**

Last Labs **Date** / /

Li = __ VPA= ___ TSH= ___

Creat= __

Current Clinical Status (check one)

__ **Depression** __ **Continued Sx**

__ **Hypomania** __ **Recovering**

__ **Mania** __ **Recovered**

__ **Mixed*** **X** **Roughening**

If new episode, estimate onset date: / /

Other Dx: _____

CGI **4** (1-7) **GAF** week **55** (0-90) **GAF** month **55** (0-90)

Comments: **28 yo MWF with BD-II – subsyndromal hypomania.**

Discussed increase LTG versus CBT and better sleep hygiene.

Plan: **Use CBT. Reassess LTG dose next visit.**

RTC **2 weeks**

Signature

FIGURE 23–1. Clinical monitoring form–representative time points (D).

A, Initial intake. B, Four weeks after starting lamotrigine. C, Regular follow-up visit during 2½ years of euthymic mood. **D**, Subsyndromal mood elevation during euthymic mood period.

esteem to more prominent and functionally disabling depressive episodes.

Approximately 6 years after the onset of these episodes, Ms. M.S. first sought psychiatric care, and she was diagnosed with "double depression" (major depressive disorder plus dysthymic disorder). Citalopram was used for 2 months without significant change in mood but with intolerable nausea and agitation. Sertraline remained ineffective after a 2-month trial and was associated with nausea and irritability. Duloxetine was effective within days of use, and the patient began to feel as if her whole life had been abnormally depressed until therapy with duloxetine. She began to see that there were many new opportunities for her life, and colors appeared brighter. After 2 months of treatment with duloxetine, however, her depression symptoms returned. The patient lost confidence in treatment and did not see a psychiatrist for another 2 years. On returning to psychiatric care, the patient was started on bupropion, which caused intolerable irritability and agitation. Due to the inefficacy of prior antidepressant medication trials, a consultation was sought.

Other Significant Findings From Assessment

During the consultation, in addition to the chronic depression overlaid with periods of more pervasive symptoms, other episodes with 2-week periods of time during which the patient was described by others as "weird" were elicited. Ms. M.S. remembered that between her more pronounced depressive episodes, she would have periods of improved mood in which she would try to catch up on her lost productivity, actually feeling that she could do anything and engaging in enough projects to more than make up for the periods of decreased productivity. During these periods, she would be more jovial. She slept 8 hours, in contrast to requiring the usual 9–10 hours, and she felt energetic during the day. Her thoughts would flow quickly, and her physical activity was increased to the level that peers commented that she was acting "weird." She would spend more and have increased libido but did not act outside her current relationships on these drives. She did not recall that these episodes were problematic at the time, until her husband reminded her of the associated irritability and that she would occasionally get into arguments with her boss and husband during these times. These hypomanic episodes were present beginning in her mid-20s but have decreased in intensity over the years.

She has described more recent periods of increased irritability with difficulty sitting still for very long and thoughts jumping between different goals she felt she needed to pursue. During these times, she sleeps a "normal" amount of 7–8 hours. These periods intersperse the long periods of depression and last only 2 to 3 days.

Ms. M.S. has no significant medical problems. She used marijuana between the ages of 15 and 18 but has not used any other illicit drugs or alcohol.

Ms. M.S. does report a family history of depression in a brother, her mother is on divalproex and paroxetine for bipolar disorder, and her maternal grandmother received electroconvulsive therapy after a postpartum depression.

Ms. M.S. had a Bipolarity Index score of 60: Episode Characteristics=10 points, Age of Onset=15 points, Course of Illness=5 points, Response to Treatment=10 points, Family History=20 points (Figure 23–2).

DSM-IV-TR Diagnosis

Axis I	Bipolar disorder, type II, current major depressive episode with rapid cycling course of illness
Axis II	None
Axis III	None
Axis IV	Moderate stressors (school, relationships, finances)
Axis V	GAF score: 50

Treatment Plan Considerations

Bipolar Depression

Based on the obtained history, supported by the Bipolarity Index score of 60, the probable diagnosis of bipolar depression was made. In addition, the rapid cycling course of illness would also direct treatment decisions.

Major depressive episodes are commonly seen in the context of a major depressive disorder, but importantly, at least one in four cases may be related to an underlying bipolar disorder (Hasin et al. 2005). Prognostic and treatment differences between a major depressive episode occurring in the context of either bipolar disorder or major depressive disorder

BIPOLARITY INDEX

For each of the items below, circle the score next to the characteristic that best describes the patient.
Characteristics' scores range from 0 (no evidence of bipolar disorder) to 20 (most convincing characteristic of bipolar disorder).

I. Episode Characteristics	
20	▪ Documented acute mania or mixed episode with prominent euphoria, grandiosity, or expansiveness and no significant general medical or known secondary etiology.
15	▪ Clear-cut acute mixed episode or dysphoric or irritable mania with no significant general medical or known secondary etiology.
(10)	▪ Clear-cut hypomania with no significant general medical or known secondary etiology. ▪ Clear-cut cyclothymia with no significant general medical or known secondary etiology. ▪ Clear-cut mania secondary to antidepressant use.
5	▪ Clear-cut hypomania secondary to antidepressant use. ▪ Episodes with characteristic sxs of hypomania, but sxs, duration, or intensity are subthreshold for hypomania or cyclothymia. ▪ A single MDE with psychotic or atypical features (Atypical is 2 of the following sxs: hypersomnia, hyperphagia, leaden paralysis of limbs) ▪ Any postpartum depression.
2	▪ Any recurrent typical unipolar major depressive disorder. ▪ History of any kind of psychotic disorder (i.e., presence of delusions, hallucinations, ideas of reference, magical thinking).
0	▪ No history of significant mood elevation, recurrent depression, or psychosis.
II. Age of Onset (1st affective episode/syndrome)	
20	▪ 15 to 19 years
(15)	▪ before age 15 or between 20 and 30
10	▪ 30 to 45 years
5	▪ after age 45
0	▪ No history of affective illness (no episodes, cyclothymia, dysthymia, or BP NOS).
III. Course of Illness / Associated Features	
20	▪ Recurrent, distinct manic episodes separated by periods of full recovery.
15	▪ Recurrent, distinct manic episodes with incomplete inter-episode recovery. ▪ Recurrent, distinct hypomanic episodes with full inter-episode recovery.
10	▪ Comorbid substance abuse. ▪ Psychotic features only during acute mood episodes. ▪ Incarceration or repeated legal offenses related to manic behavior (e.g., shoplifting, reckless driving, bankruptcy).
(5)	▪ Recurrent unipolar MDD with 3 or more major depressive episodes. ▪ Recurrent, distinct hypomanic episodes without full inter-episode recovery. ▪ Recurrent medication non-compliance. ▪ Comorbid borderline personality disorder, anxiety disorders, or eating disorders, or history of ADHD. ▪ Engagement in risky behaviors that pose a problem for patient, family, or friends. ▪ Behavioral evidence of perimenstrual exacerbation of mood symptoms.
2	▪ Baseline hyperthymic personality (when not manic or depressed. ▪ Marriage 3 or more times (including remarriage to the same individual. ▪ In two or more years,, has started a new job and changed jobs after less than a year. ▪ Has more than two advanced degrees.
0	▪ None of the above.
IV. Response to Treatment	
20	▪ Full recovery within 4 weeks of therapeutic treatment with mood stabilizing medication.
15	▪ Full recovery within 12 weeks of therapeutic treatment with mood stabilizing medication or relapse within 12 weeks of discontinuing tx. ▪ Affective switch to mania (pure or mixed) within 12 weeks of starting a new antidepressant or increasing dose.
(10)	▪ Worsening dysphoria or mixed symptoms during antidepressant treatment subthreshold for mania. ▪ Partial response to one or two mood stabilizers within 12 weeks of therapeutic treatment. ▪ Antidepressant-induced new or worsening rapid-cycling course.
5	▪ Treatment resistance: lack of response to complete trials of 3 or more antidepressants. ▪ Affective switch to mania or hypomania with antidepressant withdrawal.
2	▪ Immediate near complete response to antidepressant withdrawal.
0	▪ None of the above, or no treatment.
V. Family History	
(20)	▪ At least one first degree relative with documented bipolar illness.
15	▪ At least one second degree relative with documented bipolar illness. ▪ At least one first degree relative with documented, recurrent unipolar MDD and behavioral evidence suggesting bipolar illness.
10	▪ First degree relative with documented, recurrent unipolar MDD or schizoaffective disorder. ▪ Any relative with documented bipolar illness or recurrent unipolar MDD and behavioral evidence suggesting bipolar illness.
5	▪ First degree relative with documented substance abuse. ▪ And relative with possible bipolar illness.
2	▪ First degree relative with possible recurrent unipolar MDD. ▪ First degree relative with diagnosed related illness: anxiety disorders, eating disorders, ADD/ADHD.
0	▪ None of the above, or no family psychiatric illness.
60	**← Total score (0 – 100)**

FIGURE 23–2. Bipolarity Index.

need to be considered carefully, because treatment priorities differ between these diagnoses. Guidelines on the treatment of bipolar depression integrate both efficacy data and adverse effects considerations laying out an evidence-based hierarchy of treatment options.

The correct diagnosis of bipolar disorder often is delayed by up to 10 years (Hirschfeld et al. 2003). Bipolar depression has the same diagnostic criteria as major depressive disorder, and this may partially account for the delay in accurate diagnosis of bipolar disorder. However, some clinical factors may be discriminating (Table 23–1). Historically defined "atypical depressions" are often more typical for bipolar depression presentations (Akiskal 2005). Symptoms of bipolar depression are often the reverse of symptoms during mania. Thus, decreased need for sleep, decreased appetite, and psychomotor agitation in mania may switch to increased sleep, increased appetite and weight gain, and psychomotor retardation during the depressive phase. An earlier age of onset is more typical of bipolar disorder. In fact, nearly half of childhood (Geller et al. 2001) and nearly 50% of young adult or adolescent (Goldberg et al. 2001) severe depressions eventually develop into bipolar disorder instead of major depressive disorder. That clinical factor alone doubles the risk of bipolar outcome from one out of four to one out of every two patients with major depressive episodes. A family history of bipolar disorder in a first-degree relative is associated with at least a 20% chance of bipolar disorder in any given individual, which is considerably higher than the general population rate of 3% (Hasin et al. 2005).

TABLE 23–1. Discriminating between bipolar and unipolar depression

Bipolar more likely	Unipolar more likely
Symptoms	
Oversleeping	Insomnia
Overeating	Decreased appetite
Sluggishness	Restlessness
Delusions or hallucinations	Physical complaints
Mood swings or elevations	
Onset and course	
Onset age < 25	Onset age > 25
≥5 depressions total	Current depression > 6 months
Family history	
Bipolar disorder	No bipolar disorder

Source. Adapted from Mitchell et al. 2008.

Collaborative Care Model

Including family members or other sources of collateral information is particularly important in identifying hypomanic episodes. Oftentimes, patients do not recognize the association between mood symptoms and overperformance or underperformance. Family members are often the best source of objective historical information. In addition, family members are often the best subsequent support system for patients. Manualized family-focused therapy for bipolar disorder, in combination with pharmacotherapy, has been shown to be more effective at minimizing mood symptoms recurrence and maximizing medication compliance than usual pharmacotherapy with crisis management (Miklowitz et al. 2003). Although brief family psychoeducation is included as part of the standard initial clinical evaluation, for this case we chose manualized CBT for bipolar disorder as the psychotherapeutic adjunctive treatment. The choice of type of adjunctive psychotherapy is further discussed later.

Guideline Selection

The most recent guideline available at the time the patient was seen was the Texas Implementation of Medication Algorithm (TIMA), which integrates benefits and risks based on the most currently available clinical trials, with particular focus on double-blind, placebo-controlled trials. The last American Psychiatric Association guideline for treatment of bipolar disorders available at the time we saw this patient was published in 2002, although a contemporary revision has been published.

Currently, TIMA suggests a tiered approach to treatment of bipolar depression (Table 23–2; Suppes et al. 2005). At the first stage of treatment, lamotrigine monotherapy or lamotrigine plus an antimanic agent is suggested, based on a post hoc analysis of lamotrigine monotherapy in bipolar I depression

TABLE 23–2. Texas Implementation of Medication Algorithm for bipolar I depression

	Therapy
Stage 1	Lamotrigine + antimanic agent *or* Lamotrigine monotherapy (optimize lithium if already on lithium)
Stage 2	Quetiapine *or* olanzapine-fluoxetine combination (OFC)
Stage 3	Combination from lithium, lamotrigine, quetiapine, *or* OFC
Stage 4	Any of the following: lithium, lamotrigine, quetiapine, OFC, valproate, carbamazepine combined with any of the following: selective serotonin reuptake inhibitor, bupropion, venlafaxine *or* Electroconvulsive therapy
Stage 5	Monoamine oxidase inhibitor, tricyclic antidepressant, pramipexole, atypical antipsychotic, oxcarbazepine, inositol, stimulants, thyroid augmentation

Source. Adapted from Suppes et al. 2005.

(Calabrese et al. 1999). The combination therapy is recommended for patients with history of severe manic episodes, because lamotrigine is only modestly preventive of manic episodes (Goodwin et al. 2004) and has no acute antimanic efficacy (Bowden et al. 2000). For patients already receiving lithium therapy, optimizing the lithium dose to target a blood level of 0.8 mEq/L is also a first-line recommendation (Nemeroff et al. 2001).

At the next stage of treatment, two U.S. Food and Drug Administration–approved treatments are suggested, including olanzapine plus fluoxetine and quetiapine monotherapy, based on robust placebo-controlled trials (Calabrese et al. 2005; Thase et al. 2006; Tohen et al. 2003). Despite the faster onset of action and more robust studies, both of these medications have greater adverse effects burden and so are recommended at the second stage of treatment. A more clinically understandable and useful way to understand this balance between efficacy and risk data is to use number-needed-to-treat (NNT) and number-needed-to-harm (NNH) calculations (Ketter 2010). Thus, although olanzapine and quetiapine had NNTs for acute bipolar depression response, compared with placebo, of 4 and 6, respectively, they had NNHs for at least 7% weight gain (olanzapine) and sedation (quetiapine), compared with placebo, of 6 and 5, respectively—meaning that side effects were about as likely as was response, making these agents lower-priority treatment options (Ketter 2010).

The third stage of treatment combines therapies from the first and second stages of treatment. It is not until the fourth stage of treatment that antidepressant medications are suggested.

Use of Antidepressant Medications in Bipolar Disorder

Whereas first-line treatment of major depressive disorder is antidepressant medication, bipolar disorder treatment guidelines suggest mood stabilizer medications such as lithium, lamotrigine, and quetiapine (Suppes et al. 2005). Despite the available literature demonstrating efficacy of such mood stabilizers for bipolar depression, the community standard is still antidepressant medications. In fact, about one-half of patients with bipolar disorder are initially prescribed an antidepressant medication, whereas in stark contrast less than one-quarter were prescribed mood stabilizers (anticonvulsants 17% and lithium 8%) (Baldessarini et al. 2007). Adjunctive antidepressant medications have not been adequately studied in bipolar depression, and some studies do not suggest efficacy beyond mood stabilizer treatment alone (Nemeroff et al. 2001; Sachs et al. 2007). Antidepressant medications also have been implicated in inducing mania or accelerating mood cycle recurrences in patients with bipolar disorder. Up to two-thirds of mood episodes in patients with bipolar disorder may be antidepressant-associated mood episodes (Altshuler et al. 1995). Some studies suggest that among the classes of antidepressant medications, the dual-acting (serotonin-norepinephrine reuptake inhibitors, e.g., venlafaxine) may be worse at inducing mood cycling than either selective serotonin reuptake inhibitors or bupropion (Post et al.

2006). Coupled with these potential risks, antidepressant medication use for bipolar depression in the community may be excessive.

Cognitive-Behavioral Therapy for Bipolar Disorder

Intensive psychotherapy (CBT, Family-Focused Therapy [FFT], or Interpersonal and Social Rhythm Therapy) has been shown to be superior to a short course of psychoeducation in augmenting pharmacotherapy to stabilize bipolar disorder symptoms (Miklowitz et al. 2007). CBT with psychopharmacology has been shown to be specifically superior to medication management alone for maintaining bipolar disorder stability (Lam et al. 2003). For this case, CBT was selected for psychotherapy due to the availability of therapists trained in manualized CBT compared with the other therapies.

An important tool used in CBT for bipolar disorder is mood charting (see Figure 23–3 for example) (Denicoff et al. 1997). The mood chart is a focal point for collaboration between the patient, the psychiatrist, and the therapist. Mood is tracked prospectively so that pharmacological treatment decisions are not based on state-dependent memory. Therapists can use the mood changes and noted psychosocial stressors to more accurately recall cognition–mood–behavior interactions. Patients, as more active participants in their own care, may have greater treatment compliance.

Rapid Cycling Course of Illness

Approximately one-third of patients with bipolar disorder have a rapid cycling course of illness (Schneck et al. 2008). Documenting at least four full DSM-IV-TR mood episodes per year (at least 4 days of pervasive hypomanic symptoms and at least 2 weeks of depressive symptoms) can be challenging. Using a semistructured interview for the initial intake and subsequent clinical follow-up improves the documentation of such episodes. Oftentimes, patients can only recall "mood cycling" as a constellation of depressive and hypomanic symptoms without consideration of daily pervasiveness or meeting full duration criteria. In these situations, patients are often treated for depression as the prevailing presentation, whereas subsyndromal (by time and symptom count criteria) mixed episodes may be the relevant clinical syndrome. Antidepressant medication may worsen such clinical states, whereas mood-stabilizing medications have had more consistent evidence suggesting treatment efficacy. Mood charting may more accurately capture rapid cycling than patient recall alone.

Anxiety Symptoms

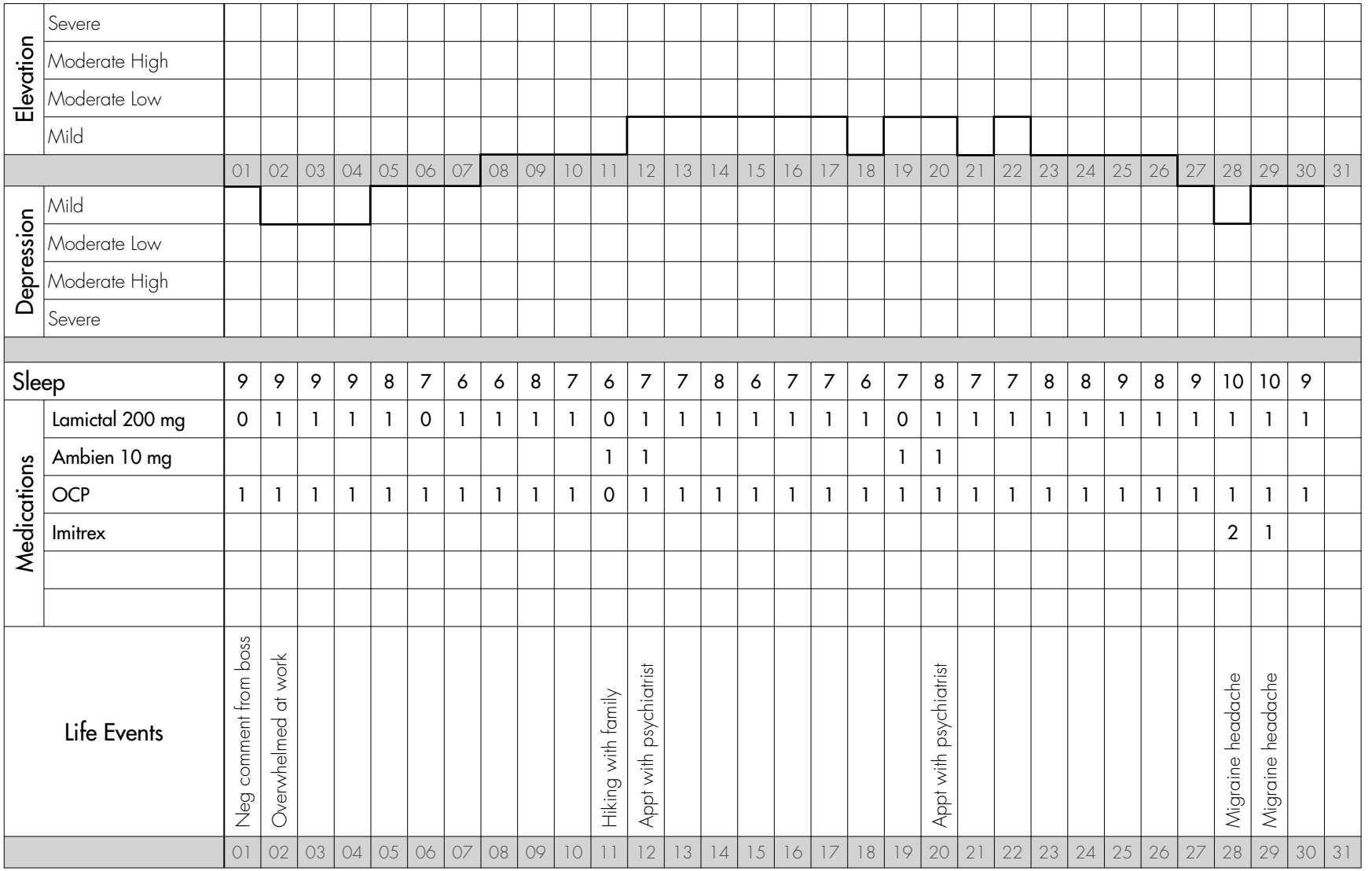
Approximately one-third of patients with bipolar disorder also have comorbid anxiety. Antidepressant medications, the mainstay of treatment for anxiety disorders, may be counterproductive for treatment of bipolar disorder; therefore, anxiety symptoms are frequently treated only if remaining after adequate mood stabilization and first targeting mood cycling and depression. CBT, on the other hand, may be used without the liability of antidepressant medications. Thus, the high comorbidity of anxiety symptoms with bipolar disorder may be another reason to preferentially use CBT as the intensive psychotherapy of choice for bipolar disorder.

Course

Ms. M.S. was started on lamotrigine, following the recommended slow titration. The risk of rash was discussed. At 200 mg/day of lamotrigine, she began to feel improvement in her mood. Mood improvement was tracked with the Clinical Monitoring Form, noting a generally steady improvement in overall symptoms even though depressed mood did not improve until approximately 4–5 weeks after starting lamotrigine. It was useful to have a standardized list of symptoms to monitor at each visit to show the patient that her mood is more than the single symptom of depressed mood or anhedonia. CBT was concurrently started and continued on a weekly basis initially. Mood charting was introduced during the initial consultation with the psychiatrist and was reinforced and discussed in depth at each CBT sessions. The psychiatrist and the therapist used the patient-generated mood chart as the focus of collaboration. Anxiety symptoms remained, but at a lower level of severity. Medication or CBT were offered as treatment options, and the patient chose CBT based on the risk-benefit discussion.

The patient remained in euthymic mood for approximately 2½ years. During that time, her visits to the psychiatrist decreased to approximately every 3 months. Her visits with her therapist continued at

Mood Chart: M.S. Month: Sept 2007



		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Sleep		9	9	9	9	8	7	6	6	8	7	6	7	7	8	6	7	7	6	7	8	7	7	8	8	9	8	9	10	10	9	
Medications	Lamictal 200 mg	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	
	Ambien 10 mg											1	1							1	1											
	OCP	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	Imitrex																												2	1		
Life Events		Neg comment from boss	Overwhelmed at work									Hiking with family	Appt with psychiatrist								Appt with psychiatrist								Migraine headache	Migraine headache		
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31

FIGURE 23–3. Mood charting—corresponds approximately to Clinical Monitoring Form, time point D.

HA=headache; OCP=oral contraceptive pill.

twice per month, then decreased to once per month. During the first year of euthymic mood, the patient had a slight, subsyndromal mood elevation around Labor Day, which was noticed only as the minor decreased sleep noted on the mood charting (Figure 23–3). The therapist asked the patient to call the psychiatrist to discuss a medication reevaluation. Reinitiating weekly CBT sessions was chosen instead of increasing the lamotrigine. Symptoms never reached DSM-IV-TR syndromal level and resolved within 3 weeks. The patient then returned to her quarterly psychiatric and monthly therapist visits.

The patient then became interested in becoming pregnant. She had tried to decrease lamotrigine dosage for the planned pregnancy, but she returned to 200 mg/day when some irritability returned. When she did become pregnant, the pregnancy was otherwise unremarkable. She returned to twice-monthly therapy and monthly psychiatric visits. Within 1 month postpartum she had a major depressive episode, despite arranging weekly therapist and twice-monthly psychiatric visits. Mood stabilized after lamotrigine was increased to 225 mg/day, and she remained stable for another 15 months until the birth of her second child. She had another major depressive episode within 1 month of delivery, but she has since been stable for 3 years after lamotrigine was increased to 300 mg/day.

Summary of Guideline Use

For this patient with bipolar depression, we used the TIMA guidelines suggesting an initial trial with either lamotrigine or lithium. Given the better tolerability profile of lamotrigine compared with lithium, we started this patient on a trial of lamotrigine, which led to improvement at the target dosage of 200 mg/day given as a single daytime dose. Although quetiapine is listed as a second-stage treatment option due to greater adverse effects burden, it may be considered as an initial treatment based on a faster onset of action. Lamotrigine is slower to onset of improvement but better tolerated, whereas quetiapine may show benefit by the first week of treatment but generally has more tolerability challenges. It should also be considered that quetiapine has benefits for anxiety symptoms, which were present in this patient. This patient was not overburdened by a severely functionally disabling illness, so lamotrigine was chosen for its greater tolerability.

Even prior to using a specific treatment guideline, for bipolar disorder it is even more imperative to make an accurate diagnosis. Bipolar disorder is an important consideration in a patient who presents with a major depressive episode, due to differences in evidence-based treatment preferences. Utilizing collateral sources of information serves dual purposes of providing more objective clinical history and establishing a collaborative care model.

Clinical monitoring of patient with bipolar disorder requires a collaborative approach involving family and peer supports, and therapists as well. TIMA advocates combined pharmacotherapy and psychotherapy but does not indicate a preference for any specific psychotherapy, owing to a lack of direct comparisons between psychotherapies. CBT is usually chosen in our setting due to the availability of therapists in our community with specialized training in CBT.

Mood charting played a pivotal role in identifying the subsyndromal mood elevation during a period in this patient's follow-up when symptoms had remitted for many months and the frequency of clinical visits with psychiatrist and therapist were decreasing. Collaboration with the therapist clearly closed the information loop for having the patient inform the psychiatrist about the mood elevation symptoms. Tracking specific symptoms at each visit with the Clinical Monitoring Form improved sensitivity for noting clinical improvement early in treatment and some mood destabilization later in ongoing clinical management.

Based on treatment guidelines, we could have either increased lamotrigine or added lithium or an atypical antipsychotic medication for this noted subsyndromal mood elevation. Instead, to control this mood recurrence, we offered these medication options or increasing psychotherapy frequency to improve sleep hygiene and medication compliance and decreasing external stimulus. Based on the collaborative care model, patient input is equally important, so the decision was made to increase psychotherapy sessions. In addition, we scheduled a follow-up visit within 2 weeks to ensure that pharmacological interventions were always available.

Ways to Improve Practice

Patients and families often would like to investigate their illness on their own. The Internet is a vast but

sometimes too unregulated resource. We have offered various consumer-oriented books on bipolar disorder, but a resource list compiled by consumers may be more effective.

CBT is an effective adjunctive treatment for patients with bipolar disorder, and it is broadly applicable to many patients. However, the main reason that FFT and Interpersonal and Social Rhythm Therapy were not used was the lack of reliable referral options. A better understanding is needed of the availability of community resources for these evidence-supported therapies.

Patients frequently benefit from peer-support groups as a further adjunct to treatment. However, most clinicians only have a basic understanding of the workings of local peer-support groups. We have started to serve as medical advisors to several local groups in order to extend the concept of collaborative care further.

Family involvement is greatest at the initiation of care due to the acuity of illness. However, over time, with longer periods of euthymic mood and prolonged normality, family involvement diminishes. We cannot help but wonder whether continuing to actively involve family and/or peer supports would have caught the subsyndromal mood elevation earlier. Involving family and peer supports throughout the acute and stable periods may somewhat destigmatize coming to the psychiatrist's office. We will consider regularly inviting family and/or peer supports to attend office visits.

References

Akiskal HS: The dark side of bipolarity: detecting bipolar depression in its pleomorphic expressions. J Affect Disord 84:107–115, 2005

Altshuler LL, Post RM, Leverich GS, et al: Antidepressant-induced mania and cycle acceleration: a controversy revisited. Am J Psychiatry 152:1130–1138, 1995

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

American Psychiatric Association: Practice guideline for the treatment of patients with bipolar disorder. Am J Psychiatry 151 (12 suppl)1–36, 2002. Available at: http://www.psychiatryonline.com/pracGuide/pracGuideTopic_8.aspx.

Baldessarini RJ, Leahy L, Arcona S, et al: Patterns of psychotropic drug prescription for U.S. patients with diagnoses of bipolar disorders. Psychiatr Serv 58:85–91, 2007

Bowden C, Calabrese J, Ascher J, et al: Spectrum of efficacy of lamotrigine in bipolar disorder: overview of double-blind placebo-controlled studies. Paper presented at the 39th annual meeting of the American College of Neuropsychopharmacology, San Juan, Puerto Rico, December 2000

Calabrese JR, Bowden CL, Sachs GS, et al: A double-blind placebo-controlled study of lamotrigine monotherapy in outpatients with bipolar I depression. Lamictal 602 Study Group. J Clin Psychiatry 60:79–88, 1999

Calabrese JR, Keck PE Jr, Macfadden W, et al: A randomized, double-blind, placebo-controlled trial of quetiapine in the treatment of bipolar I or II depression. Am J Psychiatry 162:1351–1360, 2005

Denicoff KD, Smith-Jackson EE, Disney ER, et al: Preliminary evidence of the reliability and validity of the prospective life-chart methodology (LCM-p). J Psychiatr Res 31:593–603, 1997

Derry S, Moore RA: Atypical antipsychotics in bipolar disorder: systematic review of randomised trials. BMC Psychiatry 7:40, 2007

Geller B, Zimerman B, Williams M, et al: Bipolar disorder at prospective follow-up of adults who had prepubertal major depressive disorder. Am J Psychiatry 158:125–127, 2001

Goldberg JF, Harrow M, Whiteside JE: Risk for bipolar illness in patients initially hospitalized for unipolar depression. Am J Psychiatry 158:1265–1270, 2001

Goodwin GM, Bowden CL, Calabrese JR, et al: A pooled analysis of 2 placebo-controlled 18-month trials of lamotrigine and lithium maintenance in bipolar I disorder. J Clin Psychiatry 65:432–441, 2004

Hasin DS, Goodwin RD, Stinson FS, et al: Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry 62:1097–1106, 2005

Hirschfeld RM, Calabrese JR, Weissman MM, et al: Screening for bipolar disorder in the community. J Clin Psychiatry 64:53–59, 2003

Ketter TA (ed): Handbook of Diagnosis and Treatment of Bipolar Disorders. Washington, DC, American Psychiatric Publishing, 2010

Lam DH, Watkins ER, Hayward P, et al: A randomized controlled study of cognitive therapy for relapse prevention for bipolar affective disorder: outcome of the first year. Arch Gen Psychiatry 60:145–152, 2003

Miklowitz DJ, George EL, Richards JA, et al: A randomized study of family focused psychoeducation and pharmacotherapy in the outpatient management of bipolar disorder. Arch Gen Psychiatry 60:904–912, 2003

Miklowitz DJ, Otto MW, Frank E, et al: Psychosocial treatments for bipolar depression: a 1-year randomized trial from the Systematic Treatment Enhancement Program. Arch Gen Psychiatry 64:419–426, 2007

Mitchell PB, Goodwin GM, Johnson GF, et al: Diagnostic guidelines for bipolar depression: a probabilistic approach. Bipolar Disord 10:144–152, 2008

Nemeroff CB, Evans DL, Gyulai L, et al: Double-blind, placebo-controlled comparison of imipramine and paroxetine in the treatment of bipolar depression. Am J Psychiatry 158:906–912, 2001

Post RM, Altshuler LL, Leverich GS, et al: Mood switch in bipolar depression: comparison of adjunctive venlafaxine, bupropion and sertraline. Br J Psychiatry 189:124–131, 2006

Sachs GS, Nierenberg AA, Calabrese JR, et al: Effectiveness of adjunctive antidepressant treatment for bipolar depression. N Engl J Med 356:1711–1722, 2007

Schneck CD, Miklowitz DJ, Miyahara S, et al: The prospective course of rapid-cycling bipolar disorder: findings from the STEP-BD. Am J Psychiatry 165:370–377; quiz 410, 2008

Suppes T, Dennehy EB, Hirschfeld RM, et al: The Texas implementation of medication algorithms: update to the algorithms for treatment of bipolar I disorder. J Clin Psychiatry 66:870–886, 2005

Thase ME, Macfadden W, Weisler RH, et al: Efficacy of quetiapine monotherapy in bipolar I and II depression: a double-blind, placebo-controlled study (the BOLDER II study). J Clin Psychopharmacol 26:600–609, 2006

Tohen M, Vieta E, Calabrese J, et al: Efficacy of olanzapine and olanzapine-fluoxetine combination in the treatment of bipolar I depression. Arch Gen Psychiatry 60:1079–1088, 2003

24

Obsessive-Compulsive Disorder

Bibi Das, M.D.
Lorrin M. Koran, M.D.

Setting

The patient was seen in a private practice setting. The population in the surrounding area is of high educational attainment and upper socioeconomic status and is ethnically diverse. The patient population is also quite sophisticated in accessing information on the Web.

The treatment strategy utilized is evidence-based pharmacotherapy combined with cognitive-behavioral therapy (CBT).

Illustration

This study illustrates:

- Combining medications (selective serotonin reuptake inhibitors [SSRIs]) and psychotherapy (CBT) in the treatment of obsessive-compulsive disorder (OCD)
- The importance of family involvement in the treatment to eliminate family accommodation to the symptoms
- Cognitive restructuring in an attempt to increase motivation for exposure and response prevention (ERP)
- Repeated Yale-Brown Obsessive Compulsive Scale (Y-BOCS; Goodman et al. 1989) assessments of the severity of the disorder
- The steps taken to improve adherence
- The utility of assessing common co-occurring disorders

Chief Complaint

Ms. D.B., a 19-year-old single Caucasian female, presents with a long history of "taking hours to get anything done."

Present Illness

Ms. D.B. reported that she has been "stuck in life." She is unable to get anything done "because everything I do takes forever." She has intrusive thoughts of hurting her family that make her feel "sick to the stomach." She relieves her anxiety by washing her hands or saying a prayer under her breath. Over the years, she has developed elaborate rituals that dictate how she showers, eats, drives, studies, and carries out many other activities. She spends 7–8 hours a day doing her rituals.

Ms. D.B. is afraid that she is getting "stupid and crazy because I do things that make no sense." She feels helpless in the face of her thoughts because she cannot stop herself from carrying out her rituals.

Ms. D.B. reported that she has had rituals "ever since I was playing with dolls." In the previous 5 months, however, her symptoms had worsened, so that she stopped going to school and was unable to drive or leave the house. She believes that the acute stressor was her having been accepted to a college far from home.

She also acknowledged that in the past month she had struggled with pervasive sad mood and low energy and has had difficulty enjoying herself. She

TABLE 24–1. Past medication trials

Medication	Daily dosage	Duration	Side effects	Efficacy
Fluoxetine	20 mg	4 months	Agitation	Mild
Citalopram	20 mg	3 months	Sexual	None
Olanzapine	10 mg	1 month	Weight gain	None
Clonazepam	0.5 mg, prn	with citalopram	None	Most helpful

feels "either guilty or stupid all the time" and has had feelings of worthlessness. She denies any hopelessness, suicidal ideation, and symptoms of mood elevation. Appetite and memory are normal. She has had problems falling asleep but denies any mid-cycle awakening, nightmares, early morning awakening, or daytime sedation. She also complained of problems with focus and distractibility. Ms. D.B. has had free-floating anxiety but denied any panic attacks. She has had no impulse-control problems.

Past Psychiatric History

Ms. D.B. has no history of substance abuse.

Her compulsions began in childhood, when her little brother was born. She went to the store and bought three candies even though they were four siblings. She remembers feeling "like a monster because I wanted one of them to die." As a child she could not talk about this to anyone and controlled her anxiety by creating rituals "that would neutralize the bad thoughts."

She denies any history of depression before the current episode. At age 17, she sought treatment from her pediatrician. She had tried several medications (Table 24-1) prescribed by her pediatrician with "little help." I verified her medication history by asking her pharmacy to fax me records. All medications had been refilled on time, indicating that Ms D.B. adhered to treatment. She had not had a trial of psychotherapy.

Medical History

- Euthyroid (checked in the previous year during her annual physical)
- No history of head injury, seizures, loss of consciousness, sexually transmitted disease, or any risk factor for HIV
- No known drug allergies

Family History

- Motor tics in a maternal cousin
- A maternal aunt with major depressive disorder and anxiety disorder (the specific type was unknown to the patient), who responded well to escitalopram
- A brother with generalized anxiety disorder who had been treated successfully with escitalopram 10 mg/day

DSM-IV-TR Diagnosis

Axis I Obsessive-compulsive disorder
Major depressive disorder, currently moderate severity

Treatment Plan Considerations

1. Compulsions—Y-BOCS score 16
2. Obsessions—Y-BOCS score 16
3. Depression—Hamilton Rating Scale for Depression (Ham-D; Hamilton 1960) score 24

A "gold standard" OCD severity measure is the Y-BOCS. A score of 32 (16+16) falls in the "severe" range (Goodman et al. 1989).

Secondary Points in the Plan

Psychoeducation will emphasize understanding that the cause of the disorder is unknown and that it represents a chemical problem in the brain and is not a weakness of character and cannot be willed away, has a waxing and waning course that seems influenced by stress, is treatable and responds to both medications and to CBT, and usually does not go away completely. We will emphasize the need for adherence to treatment in order to reap the benefits, and we will set a realistic expectation of recovery in terms of symptom relief and the time needed to reach this goal.

Treatment Goals

- Compulsions: Less than 30 minutes a day and not interfering with functioning
- Obsessions: Less than 30 minutes a day and not interfering with functioning.
- Depression: Remission

Treatment Goal Measures

- Y-BOCS: Score less than 8 (Simpson et al. 2006)
- Ham-D: Score less than 7

Course

Week 1

The options of using medications alone, CBT with ERP alone, or a combination of these two treatments were discussed (Simpson and Liebowitz 2005).

I (Dr. B.D.) prefer to meet my patients once a week if I am providing medication management together with CBT. The literature and expert opinions suggest that CBT sessions should be scheduled at least once weekly (March et al. 1997; Whittal et al. 2005). The number of treatment sessions, their length, and the duration of an adequate trial have not been established, but expert consensus recommends 13–20 weekly sessions "for most patients" (March et al. 1997). I share these data with my patients so that they can have a realistic expectation of when they can expect to start feeling better. Also, this helps them decide whether they are willing to make the time and the emotional and financial commitment required to fully participate in the therapy process.

Ms. D.B. was started on escitalopram 5 mg/day. She had tried two other SSRIs that had not brought significant symptom control and had caused some side effects. Although clinical trials indicate that all SSRIs are equally efficacious in treating both depression and OCD (Koran et al. 2007), I decided to start her on escitalopram because her brother had responded to this drug. The common side effects were discussed. In some patients, drug-drug interactions, particular unwanted side effects, and insurance coverage are important additional factors to consider, but in Ms. D.B.'s case these factors were not relevant.

I also asked Ms. D.B. to take 1,000 mg/day of eicosapentaenoic acid (EPA), an omega-3 fatty acid (available in fish oil capsules), in view of studies suggesting that at this dosage it may be a beneficial augmentation strategy in depressive disorders (Freeman et al. 2006), and encouraged her to add 800 μg/day of folic acid to help treat her depression, again based on studies suggesting possible benefit (Alpert et al. 2002; Coppen and Bailey 2000).

I provided Ms. D.B. with a reading list and asked her to procure *The OCD Workbook* (Hyman and Pedrick 2005) and *Brain Lock* (Schwartz 1996). Her assignment for the first week was to make an Anxiety/Exposure List and to grade the severity of anxiety aroused in different fear-provoking situations using the Subjective Units of Distress Scale (SUDS), with each item rated from 1 to 100 points.

Before initiating CBT, I explained the nature of the treatment, including its here-and-now focus, the rationale of underlying treatment procedures, and what she would be required to do.

At the start of therapy, I used the Y-BOCS symptom checklist (Goodman et al. 1989) to help her create a list of target symptoms, including obsessions, compulsions, and items or situations that are avoided because of OCD concerns. The patient ranked the listed items from least to most anxiety provoking.

In CBT consisting of ERP, patients are taught to confront feared situations and objects (i.e., exposure) and to refrain from performing rituals (i.e., response prevention). Exposures may include in vivo confrontations (e.g., setting the dials of the thermostat to the wrong number, wearing unlucky clothes) and imagining feared consequences. Exposures that provoke moderate anxiety are prescribed first, followed as quickly as tolerable by exposures of increasing difficulty. Patients must face their fears for a prolonged period without ritualizing, allowing the anxiety or discomfort to dissipate on its own ("habituation"). The goal is to weaken the connections between feared stimuli and distress and between ritualizing and relief from distress.

She was also encouraged to visit the Obsessive Compulsive Foundation Web site, www.ocfoundation.org, to learn more about OCD and to look for a local OCD support group sponsored by the foundation. Unfortunately, she was unable to join a group because of time constraints.

Week 2

Ms. D.B. came back after having taken escitalopram 5 mg/day for a week. She had tolerated the medication with no side effects but reported no symptom improvement. I increased the dosage by 5 mg to 10 mg/

TABLE 24–2. Obsessions/compulsions and subjective distress levels

Obsession or compulsion	SUDS level (0–100)
Thinking about, talking about, or seeing pirates	35
Locking the front door once, checking twice, and walking away	45
Leaving the bed in the morning without doing ritual	65
Not adjusting the car temperature or radio station to the correct number while driving	70
Not checking all kitchen appliances before going to bed	75
Not saying a prayer when parents leave the house	85
Not doing elaborate bath ritual	100

Note. SUDS=Subjective Units of Distress Scale.

day, because 5 mg/day was far lower than the dosage (20 mg/day) demonstrated quite effective in a controlled clinical trial (Stein et al. 2007). In contrast to improvement in depression, OCD improvement takes longer and often requires higher SSRI dosages (often above U.S. Food and Drug Administration–approved maximum dosages) (Koran et al. 2007). Many articles suggest waiting 8–12 weeks before increasing the SSRI dosage or considering a switch. However, I tend to increase the dosage faster for two reasons:

1. Delay in initial response often leads to noncompliance. Also, because it appears that higher dosages are more effective than lower dosages in treating OCD (Koran et al. 2007), slow titration wastes precious time before the patient can experience relief of symptoms.
2. SSRI side effects tend not to be strongly dose dependent, so once we reach remission with favorable tolerability, I slowly back off on the dosage over several months while being vigilant for returning symptoms. The goal is to obtain maximum symptom control with minimum medication.

Ms. D.B. had prepared an inventory of some of her OCD symptoms, as shown in Table 24–2. She reported that her obsessions revolved around fears that if she did not do her rituals, a member of her family would get into an accident and die or some natural calamity would befall that person. She reported that even though she knew that this made "no logical sense, I am a prisoner to my thoughts."

While I am working on ERP, I like to also work on the cognitive distortions that underlie OCD and try to teach the patient formal cognitive techniques aimed at changing these dysfunctional beliefs. Even though data comparing ERP alone to CBT alone remain inconclusive (Koran et al. 2007), many experts suggest that a combined approach is most fruitful. Dysfunctional beliefs that are common in OCD include magical thinking (e.g., a bathing ritual will keep my family safe), an inflated sense of responsibility for unwanted events, overestimation of the probability of feared events, the assumption that thoughts are morally equivalent to actions or inevitably lead to action ("thought-action fusion"), perfectionism, the belief that anxiety/discomfort will persist forever, and the need for control.

We decided that we would start with the OCD symptom that caused least anxiety and then move up the anxiety ladder. Thus we started with the patient's anxiety about "pirates."

Table 24–3 shows Ms. D.B.'s mini-list for her pirate anxiety. Because a SUDS level of 20 was not distressing enough, we started with writing the word "pirate." We worked on trying to spell the word, which Ms. D.B. did with little anxiety. Then, we wrote the word in the air, and finally, on paper. We

TABLE 24–3. Pirate anxiety and subjective distress levels

Activity	SUDS level
Thinking about pirates	10
Hearing someone talk about pirates	20
Writing the word *pirate*	25
Seeing a picture of a pirate	30
Seeing a movie with pirates	35

Note. SUDS=Subjective Units of Distress Scale.

TABLE 24–4. Response to intrusive thoughts

Intrusive thought	Response (talking back to your brain)
If I have bad thoughts, it means I am a bad person.	"It's only my overactive brain chemicals."
Maybe I will act on my bad thoughts.	"It's not me, it's my OCD."
Maybe this is not OCD.	This is more evidence that this is OCD.

obtained several pictures of pirates from the Internet, and I asked Ms. D.B. to look at them. The feared consequence of ritual prevention that cropped up was "my family will get hurt by bad guys."

We decided that Ms. D.B. would try to do the writing/looking at the picture exercise when with her family.

Week 3

Ms. D.B. was now taking 10 mg/day of escitalopram. She reported a 40% decrease in her sad mood and overall anxiety. She also reported that she was getting better-quality sleep and waking up feeling more refreshed. She had some constipation, however, for which she was taking Metamucil.

Her Ham-D score reflected the improvement in her depressive symptoms, having decreased from 23 to 12. Her Y-BOCS score remained almost the same (15 obsessions + 16 compulsions). As expected, her depressive symptoms responded earlier than her OCD symptoms.

Ms. D.B. reported that she was able to do her ERP exercise in spite of increased anxiety. She was so excited by her success that she rented *Pirates of the Caribbean* and was able to watch 30 minutes of it.

We looked at "self-talk strategies" that she could use in order to plow through the exposure. We reviewed the concept of being an "impartial observer" as outlined in *Brain Lock* (Schwartz 1996). The idea is to pay attention to one's inner dialogue and talk back to the intrusive thoughts with logical and directive statements rather than giving in to them (Table 24–4).

We decided to target the next OCD symptom, locking the front door and checking it. We made a mini-list, breaking down the locking and checking behavior (Table 24–5).

The mini-list was used to do the ERP in the following steps:

1. Choose the item that produces at least moderate anxiety by bringing it on and confronting it with an adequate level of anxiety.
2. Allow the discomfort to rise and tolerate it rather than trying to neutralize it.
3. Practice ritual prevention while doing the exposure (i.e., do not pray under your breath).
4. Repeat the task again and again till the SUDS decreases to 20 or less.

We also increased Ms. D.B.'s escitalopram dosage to 20 mg/day, the dosage demonstrated more effective in OCD (Stein et al. 2007), because she had tolerated 10 mg/day quite well.

Week 4

Ms. D.B. came back the following week reporting that she was able to cut down her checking behavior to three times that week. However, the anxiety had been

TABLE 24–5. Locking and checking behaviors and subjective distress levels

Behavior	SUDS level
Locking the front door once, checking twice, and walking away	45
Locking the front door once, checking once, and walking away	65
Locking the front door once, checking once, and walking away and staying away for 1 hour	85
Locking the front door once, checking once, and walking away and staying away for 4 hours	95
Locking the front door once, checking once, and walking away and staying away overnight	100

Note. SUDS=Subjective Units of Distress Scale.

so intense that she had had her mother check the door for her and reassure her that the door was safely locked. I learned that her family had often "helped her out" when she was unable to get started or move on.

I explained to her that this "help" would be counterproductive to her improvement. We decided to have her family come in for the next session (Renshaw et al. 2005).

Week 5

During the family meeting I explained how family/caregivers could inadvertently act in ways that are counterproductive to the patient's recovery. Participating or "enabling" is a family's way to protect the loved one in the face of suffering, thus offering immediate relief from OCD distress. I elaborated some typical ways that the family was vicariously participating in OCD behavior:

- Assisting with checking rituals
- Reassuring the patient that things are "safe"
- Trying to reason the OCD away by recounting endless reasons and facts that show how unreasonable the OCD fears are. (In the process, the family may be counterproductively neutralizing the obsessive thought, giving it credence and perpetuating the cycle.)

I asked the family to come up with examples of how they had fallen into any of these traps. We then did some role-playing regarding how the family members could be helpful by being coaches and cheerleaders for Ms. D.B. The following guidelines were established:

1. Anticipate that Ms. D.B. will notice an increase in anxiety before it finally subsides
2. Decrease participation gradually to reach complete disengagement at an agreed-upon time, which is non-negotiable
3. Expect resistance and possibly anger
4. When demands are made for reassurance, the family member should calmly remind Ms. D.B. that, "because I love you, I cannot participate in harmful OCD behavior."

Weeks 6–13

From week 6 to week 13 Ms. D.B. moved down the OCD symptom target list carrying out ERP. She continued to keep her ERP log, and her Y-BOCS scores dropped from 30 to 20. However, when we reached her bathing ritual, we hit a roadblock. At this point she was unable/unwilling to do any of the exposure. She reported severe anxiety and terrible feelings of guilt. She thought that "if the rituals are for my family's safety, I cannot let them go. I feel I am a bad daughter /sister if I challenge the rituals." We tried to use the strategy of "relabeling" as outlined in *Brain Lock* (Schwartz 1996). When relabeling, she would try to identify these thoughts as OCD thoughts and say to herself, "It's not me, it's my OCD." This would allow her to not regard the thoughts as part of her normal thinking and to get perspective on how absurd they are.

From week 11 to week 13 Ms. D.B. appeared to reach an impasse in confronting her remaining OCD symptoms (Figure 24–1). We tried to augment the effect of escitalopram by adding clonazepam 0.5 mg as needed (not to exceed 1 mg/day) for anxiety. She had a history of experiencing anxiety relief with clonazepam, so I tried it first. We could have tried other augmentation strategies, for example, atypical antipsychotics, clomipramine, or buspirone (Koran et al. 2007), but it seemed to me that she just needed a little help in doing her exposure, via relief from heightened anxiety.

Week 14

Ms. D.B. returned with worsening depressive symptoms (Figure 24–2). She reported that her inability to do the ERP was making her "feel like a failure," yet trying to do the exposure "just makes me way too guilty." She wanted a "guarantee that nothing will happen to my family if I try." She also said that, "deep down I must really want them dead if I have this OCD."

We tried to address the cognitive distortions that were getting in the way of her recovery. Prominent among them were the following:

- *Magical thinking/thought-action fusion:* If I think of a pirate, bad things will happen to my family.
- *All-or-nothing thinking:* If I do not shower properly, I am totally dirty.
- *Perfectionism:* It is intolerable until I do it perfectly.
- *Overresponsibility:* I must at all times be on alert so that I do not harm any innocent person.
- *What-if thinking:* What if I make a mistake? …it is not OCD? …Dr. Das is wrong?

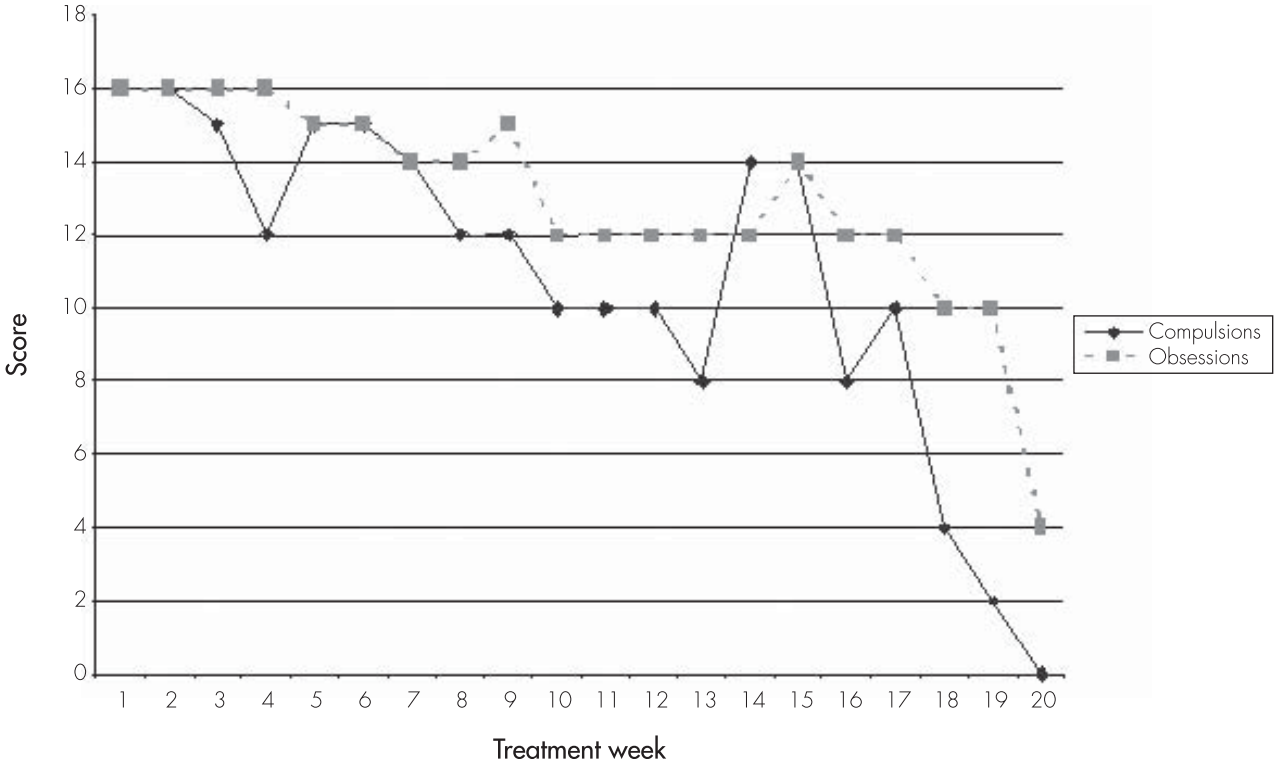


FIGURE 24–1. Patient's Yale-Brown Obsessive Compulsive Scale scores by week.

- *The exclusivity error:* All bad things will happen to my loved ones.

Ms. D.B. was to try to keep a "thought log" and to try to recognize any of these cognitive distortions if they occurred. We also decided that it would be a good idea to have a second family meeting.

Week 15

At the family meeting I had two agenda items: First, I wished to address Ms. D.B.'s guilt. I told her that maybe her OCD revolved around her family's safety because that was the dearest thing to her heart rather than the other way around. There were no data to suggest that this was a fact, but I pointed out that one can only feel intense anxiety when it involves something dear. Second, I asked all the family members if they would give Ms. D.B. permission to do her exposure. They unanimously agreed. We also agreed that if a harmful event did happen in the next few weeks, the family (and, it is hoped, Ms. D.B.) would regard it as having occurred independent of her exposure therapy.

Weeks 16–20

Ms. D.B. was able to go forward with her ERP with renewed vigor. Her Y-BOCS total score came down to 4. She continued escitalopram 20 mg/day. By week 19, she had stopped using clonazepam. We discussed the importance of medication adherence.

Ms. D.B. was to start college. I encouraged her to find a clinician to work with when she moved and gave her names of psychiatrists in the new city. I also had her sign a release of information so that I could transfer records and facilitate continuity of care.

Ways to Improve Practice

Obtain Previous Records

Obtaining records from the previous treating clinician is a good idea. In this case, Ms. D.B. knew what medications she had taken, including how long she took each one and whether it was effective. Her mother had kept records of when she had started each medication, the dosage, and how long Ms. D.B. had taken it. However, patients often do not remember such details, and time may be lost trying something that has already proved ineffective.

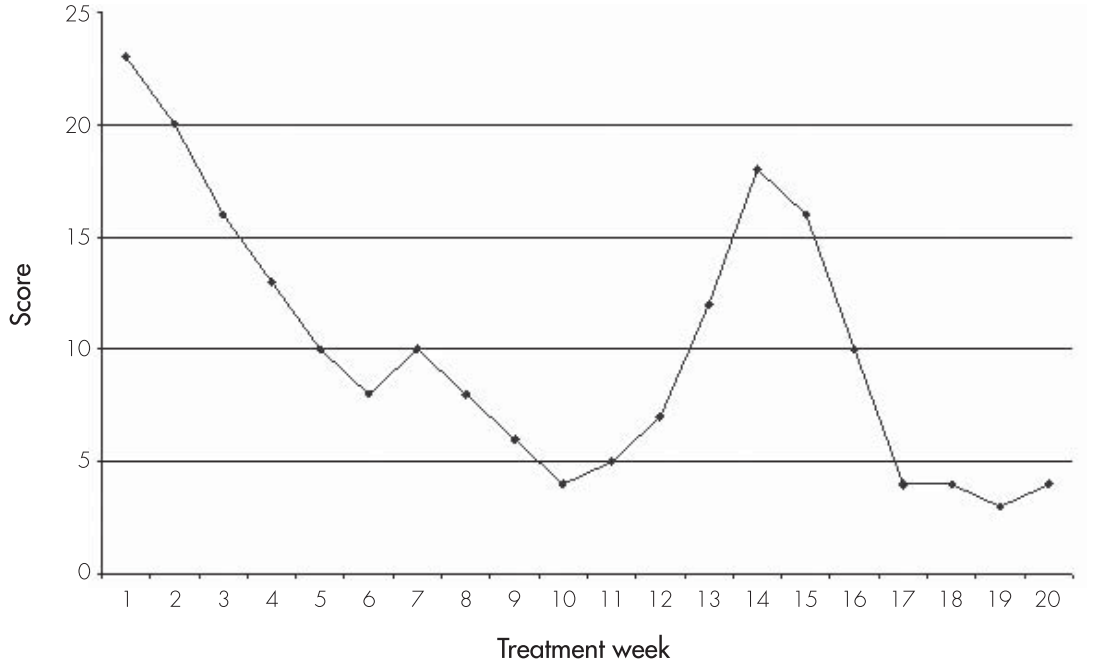


FIGURE 24–2. Patient's Hamilton Rating Scale for Depression scores by week.

Recognize Resistance

In hindsight, when I looked at Ms. D.B.'s Ham-D graph, it was obvious that she had gotten more depressed as she faced some of the challenges with ERP. The spike seen from week 6 to week 13 was due to an increase in the way she rated her "guilt." If I had paid more careful attention to the ratings and to the symptom subgroups, I could have helped her conquer her fears about doing the exposures.

References

Alpert JE, Mischoulon D, Rubenstein GE, et al: Folic acid (Leucovorin) as an adjunctive treatment for SSRI-refractory depression. Ann Clin Psychiatry 14:33–38, 2002

Coppen A, Bailey J: Enhancement of the antidepressant action of fluoxetine by folic acid: a randomized, placebo controlled trial. J Affect Disord 60:121–130, 2000

Freeman MP, Hibbeln JR, Wisner KL, et al: Omega-3 fatty acids: evidence basis for treatment and future research in psychiatry. J Clin Psychiatry 67:1954–1967, 2006

Goodman WK, Price L, Rasmussen S, et al: The Yale-Brown Obsessive Compulsive Scale (Y-BOCS), part I: development, use, and reliability. Arch Gen Psychiatry 46:1006–1011, 1989

Hamilton M: A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56–62, 1960

Hyman BM, Pedrick C: The OCD Workbook: Your Guide to Breaking Free From Obsessive Compulsive Disorder, 2nd Edition. Oakland, CA, New Harbinger Publications, 2005

Koran LM, Hanna GL, Hollander E, et al: Practice guideline for the treatment of patients with obsessive-compulsive disorder. Am J Psychiatry 164(suppl):1–53, 2007. Available at: http://www.psychiatryonline.com/pracGuide/pracGuideTopic_10.aspx. Accessed June 18, 2009

March JS, Frances A, Carpenter D, et al: The Expert Consensus Guideline Series: treatment of obsessive-compulsive disorder. J Clin Psychiatry 58(suppl):3–72, 1997

Renshaw KD, Steketee G, Chambless DL: Involving family members in the treatment of OCD. Cogn Behav Ther 34:164–175, 2005

Schwartz JM: Brain Lock: Free Yourself From Obsessive-Compulsive Behavior. New York, HarperCollins, 1996

Simpson HB, Liebowitz MR: Combining pharmacotherapy and cognitive-behavioral therapy in the treatment of OCD, in Concepts and Controversies in Obsessive-Compulsive Disorder. Edited by Abramowitz J, Houts A. New York, Springer, 2005, pp 359–376

Simpson HB, Huppert JD, Petkova E, et al: Response versus remission in obsessive-compulsive disorder. J Clin Psychiatry 67:269–276, 2006

Stein DJ, Andersen EW, Tonnoir B, et al: Escitalopram in obsessive-compulsive disorder: a randomized, placebo-controlled, paroxetine-referenced, fixed-dose, 24-week study. Curr Med Res Opin 23:701–711, 2007

Whittal ML, Thordarson DS, McLean PD: Treatment of obsessive-compulsive disorder: cognitive behavior therapy vs. exposure and response prevention. Behav Res Ther 43:1559–1576, 2005

25

Schizophrenia

A Family Psychoeducational Approach

Douglas S. Rait, Ph.D.
Ira D. Glick, M.D.

The current American Psychiatric Association (APA) practice guidelines for the treatment of schizophrenia (Lehman et al. 2004) offer a striking example of how family interventions have returned to the mainstream in contemporary psychiatric practice. Indeed, as Heru (2008) has pointed out, APA practice guidelines already recommend early couple and family involvement as well as family-based interventions for Axis I disorders, including, but not limited to, schizophrenia, bipolar disorder, major depression, panic disorder, and eating disorders. The family's role in the treatment of schizophrenia has certainly come full circle in that the very roots of the family therapy field can be found in the schizophrenia research conducted more than half a century ago by the field's founders.

These early efforts hypothesized unidirectional parent–child causal effects that, at the time, were thought to possibly explain the etiology of schizophrenia. However, the focus of family researchers in schizophrenia gradually shifted to looking at family processes that maintain the disorder (Kaplan and Rait 1993; Lehman et al. 2004; McFarlane et al. 2003). They identified a host of family factors—such as expressed emotion, stigma, isolation, and caregiver burden—implicated in the ongoing maintenance of serious psychiatric illness (Butzlaff and Hooley 1998; Kavanaugh 1992). These findings gave rise to family-centered interventions that, in tandem with psychopharmacology advances, reduce patient relapse rates and psychiatric symptoms, enhance medication compliance, and contribute to improved patient and family coping and quality of life (Cohen et al. 2008; Kaplan and Rait 1993; McFarlane et al. 2003; Pitschel-Walz et al. 2004).

Expressed emotion, perhaps the most well-known of these dimensions, is characterized by high levels of criticism, hostility, and overinvolvement on the part of family members. Research has shown that high expressed emotion is a robust predictor of relapse not only in schizophrenia but also in many psychiatric illnesses (that create caregiver burden), such as depressive disorders, bipolar disorder, and alcoholism (Hooley and Teasdale 1989; Miklowitz et al. 1988; O'Farrell et al. 1998; Rait and Glick 2008). At the same time, caregiver research has identified the high emotional and practical burdens of caring for relatives with psychiatric illnesses (Chakrabarti et al. 1992; Drapalski et al. 2008; Ferro et al. 2000).

Family psychoeducation, multiple family therapy psychoeducational groups, behavioral interventions, psychosocial rehabilitation, and self-help advocacy groups (e.g., the National Alliance on Mental Illness [NAMI]) specifically address areas of family functioning that can accelerate illness progression or,

alternately, mitigate distress and enhance coping not only for patients but also for families (Cohen et al. 2008; Dixon et al. 2009; McFarlane et al. 2003). As a result, current standards of treatment for the seriously mentally ill recognize the importance of family members' roles in the promotion of long-term recovery. Through empathic engagement, education about the illness, the expansion of social support, accessing appropriate clinical resources during periods of crisis, and the development of communication and problem-solving skills, family psychoeducation in particular has repeatedly been shown to improve both patient recovery and family well-being (Baucom et al. 1997; Glick et al. 2000; Glynn et al. 2008; McFarlane et al. 2003; Pinsof and Wynne 1995; Pitschel-Walz et al. 2004).

On the consumer front, Rolland and Walsh (2005) have also noted that both patients and family members have increasingly advocated for health care that attends to the physical and psychosocial challenges of major health conditions for all family members (not just the sickest person). Heru (2008) has suggested that "improving the family environment has important health implications equivalent to the reduction of risk factors for chronic illness" (p. 966). With this context in mind, we now look at a case example that will serve as a paradigm for the evidence-based practice in the treatment of schizophrenia. Rather than describing a single empirically validated treatment approach, *evidence-based practice* is defined as "the conscientious, explicit, and judicious use of the best evidence in making decisions about the care of individual patients" (Sackett 1999, p. 120). The assessment of and treatment decisions in this case were informed by the clinical data emerging from recent work in the area of schizophrenia, family process, and family psychoeducational interventions.

Setting

Treatment occurred in an academic medical center–based practice of psychiatry involving both brief inpatient hospitalization and long-term outpatient treatment. A flexible family systems orientation was used that combined medication management, individual and family therapy and psychoeducation, and close alliance with a consumer support network involving NAMI. The clinic's patient population covers all socioeconomic groups.

Illustration

This study illustrates the application of a flexible, evidence-based treatment of schizophrenic disorder based on APA practice guidelines. The APA practice guideline for schizophrenia emphasizes the importance of psychopharmacology combined with family intervention, during both the acute phase and the stable phase of the illness (Lehman et al. 2004):

> The acute phase is…the best time for the psychiatrist to initiate a relationship with family members, who tend to be particularly concerned about the patient's disorder, disability, and prognosis during the acute phase and during hospitalization **[I]**. Educational meetings, "survival workshops" that teach the family how to cope with schizophrenia, and referrals to local chapters of patient and family organizations such as NAMI may be helpful and are recommended **[III]**. Family members may be under considerable stress, particularly if the patient has been exhibiting dangerous or unstable behavior.… [During the stable phase:] Provide patient and family education and therapies. Work with patients to recognize early symptoms of relapse in order to prevent full-blown illness exacerbations. Educate the family about the nature of the illness and coping strategies to diminish relapses and improve quality of life for patients.

This combined treatment includes individual supportive psychotherapy, medication management, family psychoeducation, and building a network of social support for both patient and family (McFarlane et al. 2003).

Chief Complaint

The patient, Nick, is a 23-year-old single white male with a history of schizophrenia, paranoid type, and substance abuse. He was initially seen a year earlier for psychiatric follow-up (medication maintenance) several weeks after his second hospitalization. The family, also involved in his care, is middle-class and consists of Rob, the father, age 55, a sales representative; Meghan, the mother, age 52, a teacher; a son, Nick (the patient), age 23, who did not complete community college, is a part-time sales clerk, developed paranoid schizophrenia at the age of 19, and smokes marijuana; a daughter, Pam, age 20, a part-time college student who holds a full-time job; and a son, Dave, age 17, a senior in high school. All three children still live at home. They were referred for

follow-up after Nick's second hospitalization for the exacerbation of paranoid schizophrenia. Nick is currently being treated with 15 mg/day of olanzapine—a second-generation rather than a first-generation agent—although his compliance with the medication is inconsistent. He complains of side effects, primarily weight gain and sedation.

Present Illness

What Is the Current Family Problem?

Nick has both positive psychotic symptoms, characterized by auditory hallucinations, paranoid thinking, and worries about being attacked by strangers, and negative symptoms that include apathy, flat affect, and inability to seek or hold a job. He also has moderate cognitive impairment. Following his second brief inpatient stay, he was restarted on olanzapine 15 mg/day and referred back to his psychiatrist for follow-up. At the same time, family therapy was also recommended due to intensified family conflict, and the family met with a psychologist/family therapist. Stressors at home prior to that hospitalization included increased conflict around Nick's failure to show up at work, sibling conflicts, and parental conflict about whether to "force him to go to work" or allow him to stay at home and reorganize himself. Given his spotty academic and employment history, his father strongly advocated Nick's going to work "even if he didn't always like it." His mother believed this approach was counterproductive, and she urged her husband and son to "ease up on Nick" because he could not deal with the heightened stress. As a result, heated conflict between the mother and father, with increasing criticism leveled at Nick, ensued.

Why Does the Family Come for Treatment at This Time?

Following the second hospitalization, Meghan felt it was time to intensify treatment in order to develop a strategy that would address the chronic and mounting family conflict surrounding Nick's condition, the stresses undermining Nick's recovery, and the question of whether he should move out on his own. The family had already been referred by the inpatient team to NAMI's family-to-family support program, and they had received individual and family psychoeducation on the unit through lectures and group meetings.

The family reported that about the time Nick announced that he was going to move out, days after he was discharged from the hospital, the parents' longstanding quarreling intensified. The mother went to her family doctor seeking medication for her own anxiety, and she considered finding her own individual therapist because her husband "simply would not listen" to her. Upon discharge, Meghan had also tried to convene a family meeting to develop a plan.

She was unable to find a time when all family members would agree to meet ("How will talking help? He just needs to do something!"), and she requested additional family sessions with the family therapist to formulate next steps. She was joined in these sentiments by her daughter and, eventually, her younger son. However, she felt that nobody would help her and that she could not get any cooperation from the family members in supporting Nick, not to mention in doing regular household chores.

What Is the Background of the Family Problem?

Nick began to develop paranoid symptoms during high school. Rob has worked in high-tech sales for different companies over an extended period of time. Although he had been steadily employed, there were extended periods where he lost jobs, engaged in a job search, and found new work opportunities. Although this boom/bust cycle has been a common fact of life in their area, his wife felt increasingly worried about finances and the "poor role model my husband is offering to the children." Throughout this time, Meghan worked as a middle-school English teacher and managed most of the household duties. The children were all good students, although Nick was thought to have a learning disability for reading (diagnosed in middle school) that had served as an explanation for his increased apathy in class.

Family support had become more of an issue because the parents of both Rob and Meghan had had serious health problems while the children were still in the local public school. Following the death of Meghan's father due to a stroke, Meghan's mother developed dementia. The mother lived nearby and was residing in a nursing facility. Although there is no documented history of schizophrenia in the family, an uncle on Meghan's mother's side and two cousins on Meghan's father's side were similarly described as "odd, eccentric, and difficult to spend time

with." Both of Rob's parents were in poor health and lived in the Midwest near his brother and sister.

Current Interactional Patterns

The situation at home had progressively worsened during the 2 months prior to Nick's rehospitalization. Nick began to skip days at work, stay up all night, and show increased irritability. His parents engaged in intensifying conflict, with Rob accusing Nick of being the "cause for all the family's misery" while Meghan defended Nick, claiming that he was "sick," that Rob was insensitive and out of control, and that he (Rob) was the reason her level of stress was unmanageable. As the young adult children entered the conflict, Pam first joined Rob in criticizing Nick, whereas Dave loyally allied with his mother. This pattern of alliances, with Rob and Pam in coalition against Nick, Meghan, and Dave, had been pointed out to the family many times in their family sessions in the hospital. As Rob's complaints about Nick and the "indulgent" attitude shown by Meghan increased, however, Pam eventually joined Meghan in her defense, saying that although she agreed with her father's point about Nick's laziness, pot smoking, and unwillingness to take responsibility, it "isn't fair to attack Mom."

Differential Diagnosis

The patient met the criteria for schizophrenic disorder, paranoid type. In addition, his marijuana use met the criteria for substance abuse, rule out substance dependence. Finally, due to family conflict, parent–child problem was also diagnosed.

DSM-IV-TR Diagnosis

Axis I	Schizophrenic disorder, paranoid type. Substance abuse. Parent–child problem.
Axis II	None
Axis III	None
Axis IV	Deferred
Axis V	GAF score: 30. Family's score on the Global Assessment of Relational Functioning Scale: 45.

Treatment Plan Considerations

Although Nick had been treated individually for nearly 2 years without much organized family involvement (one of his parents occasionally accompanied him to monthly appointments), it was determined during his recent inpatient admission that a more consistent family intervention might be helpful in preventing relapse and future inpatient hospitalizations. The family had already been referred to NAMI, but they had never attended a meeting, nor had they attended any outpatient family support meetings offered through the clinic. Although Nick and his family accepted that schizophrenia is a brain disorder, his functional and behavioral dysfunction created ample opportunities for family members to continue to debate whether Nick was "lazy" or "sick."

Given the fact that this family had not previously participated in any organized family intervention program, developing an alliance and shared set of goals was crucial. Family members reported feeling "alone" with the problems they were experiencing, frustrated with Nick's unwillingness to take positive steps on his own behalf, and tired of the arguments and family strife. Because each family member had a busy life, with the exception of the patient, scheduling appointments where everyone could meet invariably presented a challenge. This practical challenge was addressed by setting up a loose schedule, after the initial set of intensive sessions, where meetings were scheduled monthly well in advance, with the option of meeting more frequently during periods of particular stress.

Treatment Goals, Measures, and Methods

The principal components of family psychoeducation outlined by Kaplan and Rait (1993) and McFarlane et al. (2003) include 1) coordinating all elements of treatment and rehabilitation so that everyone is working toward the same goals in a collaborative fashion; 2) paying attention to the social as well as the clinical needs of the patient; 3) providing optimal medication management; 4) assessing and supporting family strengths while acknowledging family limitations in dealing with the patient; 5) improving communication and resolving family conflict; 6) addressing feelings of grief and loss; 7) providing relevant information for family members at appropriate times; 8) developing an explicit crisis plan and response; 9) encouraging family members to expand their social systems of support to combat stigma, caregiver burnout, and social isolation; and 10) in-

TABLE 25–1. Self-reported progress and global functioning scores, by stage

Goal	Early stage	Mid-stage	Late stage
Decrease positive symptoms of disorder	x	x	x
Increase individual coping skills in work, social, and family contexts		x	x
Increase patient and family knowledge about schizophrenia	x	x	x
Reduce expressed emotion and family conflict		x	x
Normalize feelings and reduce emotional burnout	x	x	x
Reduce social isolation and increase individual and family support	x	x	x
Increase problem-solving skills and communication		x	x
Enhance patient's medication compliance	x		x
Decrease patient's marijuana abuse	x		
Global Assessment of Functioning score	30	40	50
Global Assessment of Relational Functioning score	45	55	62

creasing agency and competence and instilling hope in the face of extreme psychosocial challenges.

Specific treatment goals for Nick and his family were to increase knowledge about schizophrenia; reduce expressed emotion and family conflict; normalize family members' feelings and reduce emotional burnout; reduce social isolation and increase family support; increase family problem-solving skills and communication; encourage the patient's functioning in work, social, and family contexts; and support the patient's medication compliance.

Medication goals included decreasing positive symptoms of schizophrenia, over the long run, and improving the negative symptoms and cognition. Here, a second-generation antipsychotic, olanzapine, was used because it has somewhat better efficacy over time and causes fewer extrapyramidal effects. Olanzapine can produce weight gain, so metabolic parameters would have to be constantly monitored. In addition, given Nick's marijuana use, an additional treatment goal was to refer Nick to a substance abuse program, because comorbid substance abuse drastically lowers prognosis (Dickey et al. 2002).

In this particular case, no formal measures were used to track the problem areas or treatment goals. General progress was assessed with treatment notes in the desired areas for improvement, and estimates of progress are provided in Table 25–1 in order to show how treatment progress could be charted.

In a case in which the clinicians are using formal assessment measures to track progress, a range of rating scales and self-report instruments are available. Looking at potential measurable outcomes of family psychoeducation interventions, Cohen et al. (2008) proposed the following: greater rates of ongoing contact between family and patient's treatment team; increased empowerment for family members; increased knowledge about psychiatric illness, treatment, and resources for family members; an improved subjective recovery trajectory (such as perceived control and well-being) of the patient participating in family-based services; enhanced perceptions of family support by patients participating in family services; and improved coping and reduced stress and conflict among all involved family members. Consistent with their emphasis on evidence-based care, they also identified specific instruments to assess each of these dimensions (Cohen et al. 2008).

Course

Early Stage

In the initial family meetings, family members were introduced to the family psychoeducational approach (Heru 2008; Lehman et al. 2004). Information was shared about the etiology and various presentations of the illness, the importance of medication compliance, including the pros and cons of different medications, and the range of services and extrafamilial social supports that were available to them. Initial goals also included recognizing and

normalizing family members' emotional reactions of anger, shame, guilt, stigma, and isolation. Finally, family members learned about the structure of the family psychoeducational meetings that would serve as a format for subsequent sessions.

These psychoeducational meetings for the family followed a predictable series of steps. After a brief period of socialization and reconnecting at the session's outset, family members were asked to identify what they would like to gain from the session. Following this "go-round," a particular problem would be selected for the family to address. Then, the therapist would guide the family through a process of problem solving in which the problem would be defined and elaborated, solutions would be entertained and discussed, and a plan would then be adopted.

This process, repeated from meeting to meeting, became a reliable format that family members could use. Problems they identified included focusing on the family's pattern of alliances or "teams," the level of criticism and anger directed toward Nick, a debate over whether to return to the hospital, how to handle problems and disappointments at work, the question of whether Nick should find his own apartment, how to get Nick into a substance abuse program, ways to handle the stress of the grandmother's dementia, strategies for reducing burnout and demoralization, the issue of whether the mother should continue to teach, and how to ensure that Nick was taking his psychiatric medications. Skills for problem solving and communication were also addressed and modeled in every meeting, and the family's continuing involvement with self-help advocacy and support groups was reinforced. In collaboration with his psychiatrist, Nick complied with medication changes and had regular follow-up visits at increasingly spaced intervals. He followed a diet and exercised two to three times a week.

Individual supportive psychotherapy and medication management were started on a weekly basis during the acute phases and gradually increased to monthly sessions as Nick stabilized and his positive symptoms decreased. Family sessions were initially conducted every other week (with more frequent meetings during periods of particular crisis). After Nick's hospitalization, the family convened to discuss whether Nick should move out of the family home. Rob's disappointment and irritation because Nick had stopped going to work regularly was identified as the first problem to solve, and family members thoughtfully reexamined their roles and alliances with one another. During the session, Nick admitted that he felt he should leave because his father no longer wanted him at home. Rob reassured his son, yet he disclosed how hopeless and frustrated he was that Nick was not taking more responsibility. After all, this had been "an enormous strain for everyone." Nick's mother and siblings were able to understand Rob's feelings, and they shared their concern that Nick would deteriorate more if he lived alone and that he might smoke marijuana to excess, avoid working, and avoid taking care of himself.

Mid-Stage

Over time, Nick's positive symptoms decreased, and communication within the family improved. The family was able to find common ground when Nick agreed that he would stay at home, return to his job, and attend support groups focusing on his marijuana abuse. This concrete plan was ratified, and the family agreed to continue meeting to evaluate progress in each of the areas and to consider the longer-term question of whether Nick might find his own living situation. Feelings on the part of both siblings were addressed, and the parents reported feeling relieved that they had an "agreement that was comfortable for everyone." During periods of extreme stress and increased positive symptoms, both individual and family appointments were more frequent (as was the collaboration between the clinicians in the case). Both patient and family were engaged in treatment consisting of family psychoeducation, individual supportive psychotherapy and medication management with the psychiatrist, and social support though the family-to-family groups at the local chapter of NAMI.

Late Stage

Family meetings continued on an intermittent basis, and regular follow-up was maintained with the psychiatrist. Nick continued to refuse to deal directly with his marijuana use, and neither the family sessions nor the psychiatrist's urging resulted in his locating a treatment program that he would commit to attending. Nick's problem-solving and social skills improved. At the same time, his attendance and performance at his job improved, as he switched (because of his cognitive impairment) to a more suitable position in the stockroom instead of on the sales floor.

As Table 25–1 shows, family members reported a decrease in overall conflict, increased understanding and acceptance of Nick's illness and its consequences, and fewer feelings of frustration, anger, and demoralization. In general, the family's ability to respond to changes in Nick's behavior and affect showed greater flexibility. Family members also highlighted their ability to communicate more clearly, solve problems more effectively, and appropriately identify community and medical resources. Their final Global Assessment of Relational Functioning Scale score was 62.

Ways to Improve Practice

Although no formal measures of treatment progress were used in this case, using formal rating scales or questionnaires could provide data to both the clinicians and family members about their progress together (Cohen et al. 2008). Additionally, this flexible treatment approach combining supportive psychotherapy, medication management, family psychoeducation, and new sources of social support can be replicated in many settings, from a hospital-based practice to the private practice setting to a community mental health context. We do not recommend family therapy alone (i.e., without medication) for treatment of schizophrenia in either acute or chronic phases (Glick 2004; Kaplan and Rait 1993).

Understanding that both the biological and psychosocial aspects of the disorder and the patient's care must be recognized and respected, medical and nonmedical clinicians can work in a collaborative fashion, as is described in the case of Nick and his family, or the psychiatrist can handle both aspects of the care. In some settings that have a more developed family support program for schizophrenia and other serious mental illnesses, families can join family psychoeducational groups that serve many of the same functions as the individual family meetings described here. These groups, however, add the important dimension of social learning and additional social support that emerge when family members from different families facing similar challenges take an interest in, and literally "cross-parent" one another, while encouraging validation, problem solving, and emotional support.

Conclusions

For clinicians helping patients and families deal with the overwhelming experiences associated with schizophrenia, the proposed model of practice provides a helpful framework for an evidence-based practice that builds on considerable clinical research and satisfies current APA practice guidelines (Alexander et al. 2002; Cohen et al. 2008; Lehman et al. 2004; Patterson et al. 2004). The trajectory of schizophrenia over the patient's and family's lifetime requires a flexible treatment approach. We have described a model that combines individual and family support, psychopharmacology, education about the disorder and its consequences, skill building in the areas of family communication and problem solving, the alleviation of conflict and emotional burnout, and a recognition that appropriate involvement of both medical and extrafamilial social supports is essential.

The early work by pioneers in the field of family therapy has been updated and improved by contemporary researchers and clinicians working with patients with schizophrenia. At the same time, there is strong research support for the continued integration of family psychoeducation and other family-centered treatments with prudent psychopharmacological intervention and community support (Cohen et al. 2008; McFarlane et al. 2003). As a result of these developments, current evidence-based practices and principles of family psychoeducation, which can be administered in any clinical setting, can assist every clinician in providing optimal care to every patient and family struggling with this crippling disorder. Psychiatric training and practice would do well to emphasize the identification and application of new ways to enlist the critical involvement of family members in the care of people with schizophrenia.

References

Alexander JF, Sexton TL, Robbins MS: The developmental status of family therapy in family psychology intervention sciences, in Family Psychology: Science-Based Interventions. Edited by Liddle HA, Santisteban DA, Levant RF, et al. Washington, DC, American Psychiatric Association, 2002, pp 17–40

Baucom D, Shoham V, Muester K, et al: Empirically supported couples and family interventions for marital distress and adult mental health problems. J Consult Clin Psychol 66:53–88, 1998

Butzlaff R, Hooley J: Expressed emotion and psychiatric relapse. Arch Gen Psychiatry 55:547–552, 1998

Chakrabarti S, Kulhara P, Verma S: Extent and determinants of burden among families of patients with affective disorders. Acta Psychiatr Scand 86:247–252, 1992

Cohen AM, Glynn SM, Murray-Swank AB, et al: The family forum: directions for the implementation of family psychoeducation for severe mental illness. Psychiatr Serv 59:40–48, 2009

Dickey B, Normand SL, Weiss R, et al: Medical morbidity, mental illness, and substance use disorders. Psychiatr Serv 53:861–867, 2002

Dixon L, Goldberg R, Iannone I, et al: Use of a critical time intervention to promote continuity of care after psychiatric inpatient hospitalization. Psychiatr Serv 60:451–458, 2009

Drapalksi A, Marshall T, Seybolt D, et al: Unmet needs of families of adults with mental illness and preferences regarding family services. Psychiatr Serv 59:655–662, 2008

Ferro T, Verdeli H, Pierre F, et al: Screening for depression in mothers bringing their offspring for evaluation or treatment of depression. Am J Psychiatry 157:375–379, 2000

Glick ID: Adding psychotherapy to pharmacotherapy: data, benefits, and guidelines for integration. Am J Psychother 58:186–208, 2004

Glick I, Berman E, Clarkin J, et al: Marital and Family Therapy, 4th Edition, Washington, DC, American Psychiatric Press, 2000, pp 647–648

Glynn SM, Dixon LB, Cohen AM, et al: The family member provider outreach program. Psychiatr Serv 59:934, 2008

Heru A: Family psychiatry: from research to practice. Am J Psychiatry 163:962–968, 2008

Hooley J, Teasdale J: Predictors of relapse in unipolar depressives: expressed emotion, marital distress and perceived criticism. J Abnorm Psychol 3:229–235, 1989

Kaplan F, Rait D: Partnerships for schizophrenia: reconciling professional models. Workshop presented at the 51st annual conference of the American Association of Marriage and Family Therapy, Anaheim, CA, October 1993

Kavanaugh D: Recent developments in expressed emotion and schizophrenia. Br J Psychiatry 160:601–620, 1992

Lehman AF, Lieberman JA, Dixon LB, et al: American Psychiatric Association Steering Committee on Practice Guidelines: practice guideline for the treatment of patients with schizophrenia, 2nd ed. Am J Psychiatry 161(suppl):1–56, 2004

McFarlane W, Dixon L, Lukens E, et al: Family psychoeducation and schizophrenia: a review of the literature. J Marital Fam Ther 29:223–246, 2003

Miklowitz D, Goldstein M, Nuechterlein K, et al: Family factors and the course of bipolar affective disorder. Arch Gen Psychiatry 45:225–231, 1988

O'Farrell T, Hooley J, Fals-Stewart W, et al: Expressed emotion and relapse in alcoholic patients. J Consult Clin Psychol 66:744–752, 1998

Patterson JE, Miller RB, Carnes S, et al: Evidence based practice for marriage and family therapists. J Marital Fam Ther 30:183–195, 2004

Pinsof W, Wynne L: The efficacy of marital and family therapy: an empirical overview, conclusions, and recommendations. J Marital Fam Ther 21:585–613, 1995

Pitschel-Walz G, Leucht S, Bäuml J, et al: The effect of family interventions on relapse and rehospitalization in schizophrenia: a meta-analysis. Focus 2:78–94, 2004

Rait D, Glick I: Reintegrating family therapy training in psychiatric residency programs: making the case. Acad Psychiatry 32:76–80, 2008

Rolland J, Walsh F: Systemic training for healthcare professionals: The Chicago Center for Family Health approach. Fam Process 44:283–301, 2005

Sackett DL: How evidence-based medicine will change the way you practice. Med Econ 76:120–128, 1999

26

Substance Use Disorder Presenting as a Mood Disorder

Anna Lembke, M.D.
Keith Humphreys, Ph.D.

Setting

We practice at a university medical center with no specialty treatment program for substance use disorders. Our patients with addictive disorders are typically seeking treatment for psychiatric and/or medical issues rather than openly presenting complaints related to substance abuse.

General Treatment Approach

Our treatment approach is targeted at helping patients with addictive disorders understand that substance use is a problem and likely a contributor to their psychiatric symptoms. The challenge is to invite the patient into an honest discussion of substance use while maintaining a therapeutic alliance and addressing the chief complaint.

Illustration

This study illustrates:

- Screening for substance use disorders, specifically focusing on how to use the Timeline Followback (TLFB) method
- Differentiating primary from secondary (i.e., substance-induced) psychiatric symptoms
- Treating substance use disorders using the Veterans Health Administration/Department of Defense (VHA/DoD) *Clinical Practice Guidelines for the Management of Substance Use Disorders*

Chief Complaint

Ms. R.D., a 19-year-old female freshman undergraduate, has been referred to our outpatient psychiatric treatment clinic from the campus health center for assessment and treatment of "mood swings, insomnia, and anxiety" lasting for at least 2 months.

Present Illness

Ms. R.D. reports feeling "very sad" for 3 weeks prior to her initial presentation in our clinic, with bouts of crying, difficulty completing her schoolwork, and thoughts of suicide but no specific plan to take her life. She also reports excessive sleep, often not getting out of bed until midday, and episodes of severe anxiety, during which she begins to hyperventilate and "feel overwhelmed." She also describes a pattern of mood fluctuation, in which for up to 5 days at a time she feels happy, even euphoric, with high energy, extreme garrulousness, and decreased need for sleep—i.e., not feeling tired even if she only sleeps 3 hours the night before. She also describes some high-risk behaviors at these times: excessive spending and unprotected sex with multiple partners. She

Keith Humphreys was supported by grants from the Veterans Affairs Health Services Research and Development Service and the Robert Wood Johnson Foundation.

has not experienced an elevated mood in the past 3 weeks.

Based on her self-report narrative, Ms. R.D. appears to meet full criteria for bipolar disorder, characterized by periods of depression alternating with periods of mania/hypomania. Her mood at the time of presentation is depressed for 3 weeks, with thoughts of suicide but no specific intent or plan. Her recent behavior has put her at risk for sexually transmitted diseases.

Ms. R.D. does not mention substances in her chief complaint or opening narrative. This is not atypical, even when substances are a prominent problem in the patient's life. The reasons for this omission are manifold, having to do with but not limited to clinical inattention, stigma, and unwillingness to surrender the rewarding aspects of substance use. Therefore it falls upon the treatment provider to screen for substance use, a task that is all too frequently neglected by psychiatrists (Danielsson et al. 1999; Himelhoch and Daumit 2003).

Other Significant Findings From Assessment

Screening for Substance Use Disorders

With all of our patients, we introduce substance use screening early in the evaluation process and document our findings in a separate section just after the "History of Present Illness." We usually label it "Substance Use Patterns and History." We prioritize substance use screening for several reasons. Epidemiological studies, corroborated by our clinical experiences, tell us that 30%–40% of individuals with a lifetime mood disorder will meet criteria for a substance use disorder (Kessler et al. 1996). Likewise, many individuals dependent on nicotine, alcohol, or illegal drugs will also have a depressive disorder (Kessler et al. 2003). If substance use is detected as a problem early on, this discovery informs the rest of the interview process. It becomes necessary to delineate what relationship the presenting psychiatric symptoms have with the substance use. Substance use can mimic many psychiatric symptoms and disorders.

Many individuals, addicted or not, are poor at recalling substance use accurately, even if they are comfortable disclosing it candidly to the clinician. Therefore, asking about specific amounts and specific days using the Timeline Followback method is, in our opinion, a very effective way to begin assessment. The TLFB method charts the amounts and patterns of substance use in the preceding 1–2 weeks, independent of consequences or compulsivity of use (Sobell and Sobell 1995). This method often results in a more valid assessment of actual substance use than if the interviewer simply asks: "How much do you drink in a given week?" By keeping to specifics and a timeline, there is a markedly decreased chance of the patient minimizing consumption, mixing up drinking occasions, or mentally averaging drinking over the period of interest.

We ask the patient to begin with the day prior to our encounter and move backward, remembering how much alcohol he or she consumed on each day in the preceding 1–2 weeks. Where appropriate, we offer concrete events to stimulate memory, for example, "Let's start with Monday, that was the first day of the week before midterm exams." When asking about alcohol, it is vital to ask about the type of alcoholic beverage consumed and the size of the drink (see Table 26–1 for definitions). By knowing the type and exact amounts, we can calculate the number of standard drinks consumed in a week. Hazardous drinking for men is 14 or more standard drinks in a week or 5 or more standard drinks in a sitting. Hazardous drinking for women is 7 or more standard drinks in a week or 4 or more in a sitting (see Table 26–1). The TLFB method was developed and validated primarily for alcohol, but we also find it a useful means for other substances. It is important to remember when screening for other substances that there are no unequivocal quantity or frequency risk thresholds for hazardous use, and nearly all daily nicotine users are nicotine dependent (Veterans Health Administration/Department of Defense 2001).

Using the TLFB method with Ms. R.D., we assess amounts and patterns of alcohol use:

Interviewer: Do you drink alcohol?
R.D.: Yes.
Interviewer: Today is Wednesday. How much, if anything, did you drink last night?
R.D.: Nothing.
Interviewer: How about Monday night?
R.D.: Nothing.
Interviewer: And Sunday night—that was when the Oscars were on.
R.D.: Yes, I always watch. I didn't drink anything.

TABLE 26–1. Hazardous alcohol use screening

Definition	Comments	Male	Female
Typical drinks *per week* (U.S. Preventive Services Task Force 1996)	Standard drinks: 0.5 fluid ounces of absolute alcohol 12 ounces of beer 5 ounces of wine 1.5 ounces of 80-proof spirits	≥14	≥7
Maximum drinks *per occasion* (SAMHSA 2008)	May vary depending on age, ethnicity, medical and psychiatric comorbidity, pregnancy, and other risk factors	≥5	≥4

Interviewer: O.K., How about Saturday night? That was the night after the presidential debate.
R.D.: Well, my sorority gave a party, so of course I drank a lot.
Interviewer: How much?
R.D.: Oh, I don't know. Maybe four or five drinks.
Interviewer: What kinds of drinks did you have?
R.D.: Well, I started with a couple glasses of wine, and then I had some tequila shots. And then some vodka shots.
Interviewer: About how many ounces do you think were in each glass of wine?
R.D.: An average size glass.
Interviewer: About 5 ounces in each glass? *(gesturing with hand to mark size)*
R.D.: That looks about right.
Interviewer: And how many ounces in the tequila and vodka shots?
R.D.: I guess about an ounce.
Interviewer: And how many shots total?
R.D.: Well…maybe five…or six.

By using this method, we discover that Ms. R.D. drank eight standard drinks on the preceding Saturday night, as opposed to the "four or five" she originally quoted. By continuing the TLFB method for the remainder of the week, we discover that Ms. R.D. had consumed 18 standard drinks in the week preceding the interview. Even she is surprised by the actual amounts once it is all added up. Clearly her level of consumption is in the hazardous range (see Table 26–1) and alerts us (based on amounts and pattern of use alone) to the probability of a substance use disorder. With that information we put more time and emphasis on eliciting a detailed substance use history, now focused on DSM-IV-TR (American Psychiatric Association 2000) criteria for abuse and dependence, which primarily have to do with consequences of substance use and the compulsive drive to use.

Substance Use Patterns and History

Ms. R.D. first began using substances at age 15, when she started smoking marijuana two or three times per month. At age 17 she began to drink alcohol heavily and at age 18 began snorting cocaine. For the past 4 or 5 months, she reports drinking much more than usual for her, drinking on average three times a week, typically consuming between five and seven drinks in a single sitting. As the TLFB method revealed, she had consumed 18 standard drinks in the week preceding the interview. She reports drinking even more if there are parties in the dormitories.

She has tried to cut back on her own in the preceding couple of months but has been unable to do so. She also reports that once she begins drinking, even if she only intends to have one or two, she finds she is not able to stop drinking. She admits she plans her entire day, and even her week, around going out to drink and that it is the only thing she looks forward to all day long. She states that she lives in a sorority and that her entire social life revolves around going out and "partying" with her friends.

When asked about other drugs currently used, she endorses habitual cocaine use for 3 months, up until 3 weeks ago, when she "stopped cold turkey." Prior to 3 weeks ago, she "snorted one to two lines" during the week, and up to four or five lines per night on the weekends. She smokes approximately five cigarettes a day but can smoke up to a pack in a single sitting if she has been drinking heavily: "Then I smoke until my throat hurts."

Differentiating Primary From Secondary Psychiatric Symptoms

After gathering a detailed substance use history, we obtain a detailed mood history with the intention of sorting out the relationship, if any, between the psychiatric symptoms and the substance use. This method creates a more objective view of the relationship between substance use and psychiatric symptoms than if we asked the patient "Which came first, the depression or the drinking?" In our experience, a question asked in that way will more likely than not get a response that attributes substance use to the psychiatric illness (the "self-medication hypothesis"), even when the actual pattern is not consistent with this explanation. This is not always because patients are intentionally duplicitous, but rather because the human brain is seeking to ascribe a rational cause to an irrational behavior (i.e., substance abuse) and because framing substance use as a side effect of illness may alleviate shame in the patient's mind.

In obtaining this history, we are particularly interested in periods of sustained sobriety when we might analyze the patient's mood patterns free of the cycle of intoxication and withdrawal. We also focus on time of onset, although both mood disorders and substance use disorders tend to progress insidiously, making differentiation between primary and secondary mood disorders difficult (Mueser et al. 1998). Due to retrospective bias, clarification of primary versus secondary psychiatric symptoms is best done with prospective charting of mood during a period of confirmed abstinence.

Ms. R.D. reports that her mood swings started in high school, fluctuating wildly up and down, but that she never sought or received professional treatment for this problem and that it never impaired her functioning. She states that her mood problems predated her substance use problems and that she began drinking to "self-medicate" her depression, although this information is at odds with the substance use history she gave earlier in the interview, in which she stated that alcohol and marijuana use began at age 15, around the same time her mood problems began. When asked to clarify, she reports that she would go for a month at a time without drinking or using marijuana in high school and still had mood problems, although she has had no period of abstinence longer than a month since matriculating high school.

Focusing on more recent events, Ms. R.D. describes extreme euphoric moods alternating with depression, her primary reason for presenting to our clinic. Her high moods generally coincide with cocaine intoxication, and the low moods with cocaine withdrawal. She has been persistently depressed in the 3 weeks prior to presenting in our clinic, after her abrupt discontinuation of cocaine. Before beginning heavier cocaine use 3–4 months ago, her mood would fluctuate, but not to such an extreme.

With this information, we are now ready to make a preliminary diagnosis.

DSM-IV-TR Diagnosis

Axis I	Substance-induced mood disorder; rule out primary mood disorder (major depressive disorder vs. bipolar disorder); alcohol dependence, binge type; nicotine dependence; cocaine dependence; marijuana abuse; cocaine withdrawal
Axis II	Deferred
Axis III	At risk for sexually transmitted diseases
Axis IV	Compromised role functioning, trouble with schoolwork
Axis V	GAF scale score: 50, moderate to severe, with role impairment and passive suicidal ideation

Differential Diagnosis

We do not make a diagnosis of mood disorder at this time, because we feel based on her history and presenting symptoms that all of Ms. R.D.'s problems may be substance induced. Her euphoric moods may be entirely explained by cocaine intoxication, and her recent persistent depression by cocaine withdrawal. Her more remote history of mood fluctuations is so intertwined with substance use, without any clear antecedent or sustained period of sobriety (greater than 1 month), that we are unable to differentiate substance-induced mood fluctuations from mood disorder. Ms. R.D.'s recent anxiety and insomnia are certainly also attributable to her ongoing hazardous alcohol use and daily cigarette smoking. Nonetheless, we are not ruling out the possibility that she has an independent mood disorder, and we will use prospective analysis during a period of sobriety to make a more definitive conclusion.

Treatment Plan Considerations

Developing a specific treatment guideline for substance use disorders is difficult, because although addiction can broadly be conceptualized as one disorder, treatment interventions vary greatly depending on the types of substances being abused, the severity of the substance use disorder, an individual's readiness for change, the treatment setting, and the presence of co-occurring medical and psychiatric disorders. Of necessity then, guidelines for substance use disorders tend to be vague. Nonetheless, they are useful in promoting evidence-based strategies, particularly if they delineate not only the substance being abused but also the other parameters described earlier that influence patient care, such as readiness for change.

The American Psychiatric Association guidelines contain many useful lessons but are written at a general level, which does not substantially inform the care process step by step. In contrast, the VHA/DoD *Clinical Practice Guideline for the Management of Substance Use Disorders in the Primary Care Setting* (Veterans Health Administration/Department of Defense 2001) was most useful. [This guideline has since been replaced; see Veterans Health Administration/Department of Defense 2009.] This VHA/DoD treatment guideline is organized as an algorithm with five modules, and recommendations vary depending on treatment setting, severity of substance use, and readiness for change. The guideline focuses on assessment and management in a primary care setting. At first it may seem odd to use a primary care guideline in a psychiatry setting, but our practice is similar to primary medical practice in that our patients present for help with self-identified problems that typically have nothing to do with addiction, and those with substance abuse tend to have somewhat less severe problems than would typical be encountered in a specialty setting, for example, a substance abuse inpatient or residential program. We have outlined the VHA/DoD treatment guideline recommendations in Table 26–2, following the algorithm as it applies to our patient, Ms. R.D. In the following sections, we provide a brief review of the treatment course of Ms. R.D., comparing and contrasting our intervention with those recommended by the VHA/DoD guideline. Missing in most guidelines is the "how" to achieve these goals, and so we present in a detailed manner the exact language we use in communicating with our patients and the specific strategies we apply to achieving each clinical intervention.

Recommendation #1

- Give feedback about screening results, relating the risks of negative health effects to the patient's presenting health concerns.

In giving feedback to Ms. R.D. about her substance use, we make a conscious effort not to replicate the nagging insistence of exhausted, worried, and overwhelmed spouses/parents/friends. Instead, we take a matter-of-fact approach, emphasizing that we are partners in this endeavor of getting her well and that she is an autonomous adult who will have to make her own decisions based on the information that we provide.

Our first step is to talk about the results of the TLFB method regarding her alcohol consumption and ask what she thinks about the amount of alcohol she consumes in a week. Not surprisingly, Ms. R.D. says that she drinks no more than her friends. It is helpful to introduce the scientific finding that individuals who drink heavily tend to select in to social networks where heavy drinking is common. Thus, although she may consume the same amount as her friends, this makes it neither healthy nor by any means "normal." In fact, most young adults her age are not using substances at the same frequency and amount she does. Ms. R.D. is genuinely surprised to learn that, for example, the average 19-year-old consumes a mere fraction of the amount of alcohol she is consuming. Based on the National Survey on Drug Use and Health, 45% of women in the age range of 18–25 do not drink any alcohol in a typical week, and only 2% of 18- to 25-year-old women drink 15 or more drinks in a typical week (see Figure 26–1; Substance Abuse and Mental Health Services Administration 2008). With this information, Ms. R.D. can see that her alcohol consumption is well outside the "normal" range for women her age.

After discussing the results of the TLFB method and discussing "normal" alcohol use for a woman her age, we read aloud the diagnostic criteria for substance abuse and dependence from DSM-IV-TR. We ask Ms. R.D. if she thinks any of the criteria apply to her. She is able to endorse the compulsivity

TABLE 26–2. VHA/DoD clinical practice guideline for the management of substance use disorders for patient R.D.

Area	Recommendation	Followed
Initial assessment	Obtain history, physical examination, laboratory tests, mental status examination, and medication.	√
	Is patient medically or psychiatrically unstable or acutely intoxicated?	√ Passive suicidality present but no plan or intent: Assessed with CIWAR and not at risk for dangerous alcohol withdrawal
	Does patient exhibit hazardous substance use or abuse or dependence or risk of relapse? If yes, then…	√ Yes
Provide brief intervention	Give feedback about screening results, relating the risks of negative health effects to the patient's presenting health concerns.	√
	Inform the patient about safe consumption limits and offer advice about change.	√
	Offer to involve family members in this process to educate them and solicit their input (consent is required).	No, but took an ecological approach
	Assess patient's degree of readiness for change (e.g., "How willing are you to consider reducing your use at this time?").	√
	Negotiate goals and strategies for change.	√
	If indicated, treat with pharmacotherapy for addiction.	No, and perhaps should have considered this
	Schedule an initial follow-up appointment in 2–4 weeks.	Scheduled a follow-up in 1 week after first visit, given suicidal ideation at first presentation
	Monitor changes at follow-up visits by asking patient about use, health effects, and barriers to change.	√
	Is patient nicotine dependent? If yes, then treat.	√
	Treat concurrent psychiatric disorders, including concurrent pharmacotherapy.	√
Monitoring	Establish a specific system for monitoring substance use.	Not done, an error on our part

Note. CIWAR=Clinical Institute Withdrawal Assessment of Alcohol; VHA/DoD=Veterans Health Administration/Department of Defense.

Source. Veterans Health Administration/Department of Defense 2009.

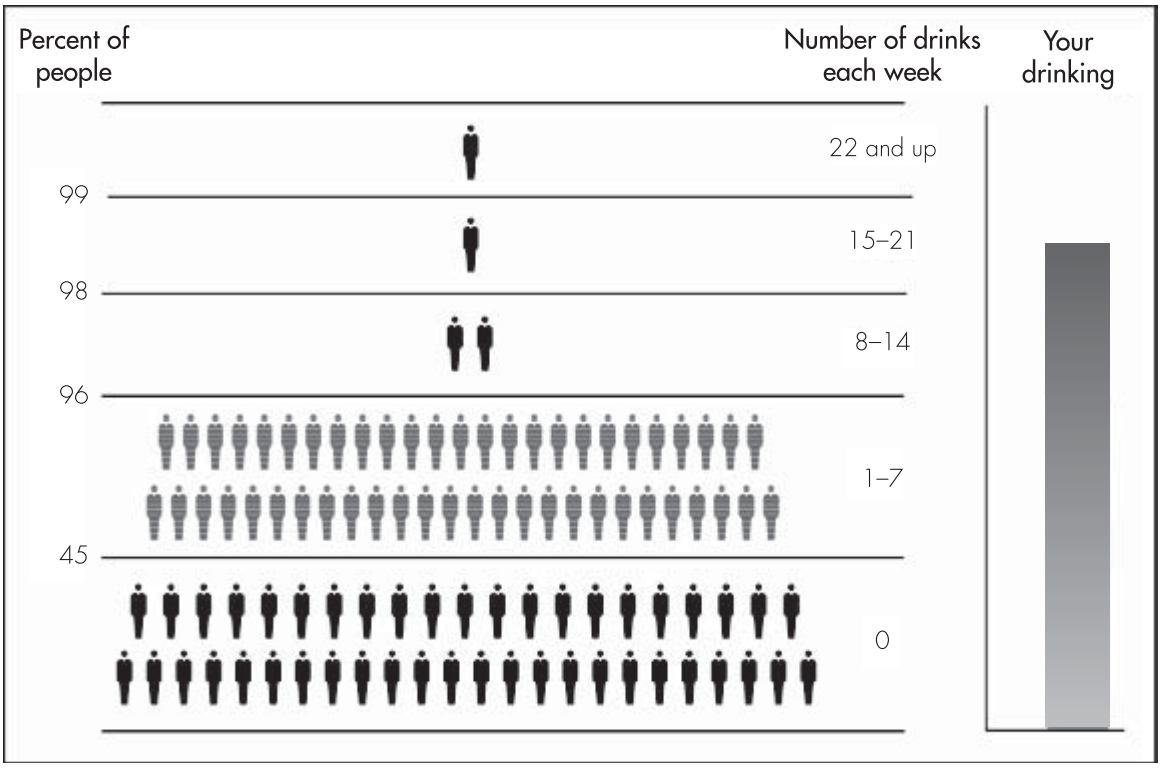


FIGURE 26–1. Rates of alcohol consumption, based on standard drinks, for women in the United States ages 18–25 years and for patient R.D.

Source. Kenneth R. Weingardt, Center for Health Care Evaluation, VA Palo Alto Health Care System and Stanford University School of Medicine; National Survey on Drug Use and Health (SAMHSA 2008).

and out-of-control nature of her use but is not able to clearly state any negative consequences. We help her identify some of the negative consequences, including exposing herself to sexually transmitted diseases; doing poorly at school; and possibly causing depression, anxiety, and insomnia, in the short-term and liver and lung damage in the long term. We talk about the cognitive and emotional problems that can result as a direct response to substance use. When the patient reports that substances "feel good" in the short term, no effort is made to argue. Following the precepts of the evidence-based treatment known as motivational interviewing, we do not try to argue patients out of what they perceive as the positives of their behavior. Rather, we attempt to balance these immediate rewards with information about longer-term costs. Ms. R.D. is able to tolerate hearing this information, but she responds with "I really don't think I am an alcoholic. If I really wanted to, I could stop drinking." Here is our opportunity then, to ask her to do just that (see Recommendation #2). Again following motivational interviewing principles, in suggesting an abstinence trial we do not label her as "alcoholic." Rather, we ask her to conduct an experiment, so to speak, to help her and us understand how serious her drinking is.

We revisit Ms. R.D.'s specific chief complaint: "mood swings, insomnia, and anxiety" and feeling "very sad" in the 3 weeks prior to her appointment. We reiterate that dysphoria, insomnia, and anxiety are the most common and pervasive symptoms of withdrawal from any addictive substance, no matter the type, and although substances may help with these symptoms in the short term, they will create these symptoms once the individual is in the cycle of intoxication and withdrawal. Many patients presenting to our clinic insist that their depression, insomnia, anxiety, and so on cause them to use substances and that if we would only "cure" their psychiatric problems, their substance use would disappear. We know based on evidence in the literature and our own clinical experience that treating the mood or anxiety disorders seldom if ever resolves the substance use disorder and that psychiatric symptoms

are in general very resistant to treatment in the context of a co-occurring substance use disorder. Numerous studies treating mood problems prospectively in patients with co-occurring substance use problems demonstrate that substance use problems do not generally resolve when mood issues are treated (Nunes and Levin 2004). Addiction appears to have an illness course independent of mood outcomes and needs its own treatment strategy. For an excellent review addressing the question of which came first, the substance use or the mood problems, please see "Dual Diagnosis: A Review of Etiological Theories," by Mueser et al. (1998). It is our very important job to educate patients that even if relief from negative emotions is what causes them to initiate substance use, the pattern of abuse and dependence is a disorder distinctly its own and needs and deserves to be treated as well.

Recommendation #2

- Inform the patient about safe consumption limits and offer advice about change.

We inform Ms. R.D. that hazardous alcohol use for an adult woman is more than seven drinks in a single week, and more than four drinks in a single sitting, and that all-cause morbidity and mortality is lowest for women who consume fewer than two standard drinks per week (Swift 1999). However, in the context of out-of-control and compulsive use, abstinence from all substances may be the best alternative, primarily because individuals with alcohol dependence have great difficulty limiting their use. Abstinence is also a recommended course of action when psychiatric symptoms are prominent, in order to help clarify what fraction of the patient's symptoms is substance induced and what fraction represents another Axis I disorder, such as major depression or panic disorder. We tell Ms. R.D. that smoking even a single cigarette a day adversely impacts her lifetime morbidity and mortality and possibly contributes to her symptoms of depression and anxiety.

After challenging Ms. R.D.'s misperception about normal substance use, discussing the criteria for a substance use disorder, educating her about the adverse consequences of substance abuse, and offering information about safe consumption, we make our first suggestions for change. For outpatients like Ms. R.D. who do not self-identify as having a substance use disorder and who present for treatment of psychiatric problems other than substance use, we typically begin by suggesting an agreed-upon period of abstinence as a test of the patient's control over substances and as a means by which to better evaluate their presenting psychiatric problems. We also let them know that even if they are unable to adhere to the contract, they may want to consider the possibility that they do not have the control they thought they had, in which case we might need a more specific and extended intervention for a substance use disorder, such as Alcoholics Anonymous. Practice guidelines recommend referral of all serious cases of drinking problems to Alcoholics Anonymous or other self-help groups, and both research literature and our clinical experience support this suggestion (Humphreys 2004).

Less commonly appreciated in cases like that of Ms. R.D., self-help groups can actually benefit patients who never attend them, because for the average person, needing to go to Alcoholics Anonymous meetings is a common-sense sign that their alcohol problem is indeed quite serious. This may be quite motivating during a trial of abstinence (e.g., "I would rather quit on my own than have to admit that I need Alcoholics Anonymous").

We also ask our patients to consider stopping smoking. Psychiatrists tend to defer smoking cessation treatment to a time when the patient is not in acute crisis. However, there is mounting evidence that depressed smokers are both willing and able to quit smoking at rates similar to nondepressed smokers and that the risk of depression in the wake of smoking cessation is much less than previously believed. Furthermore, stopping smoking may improve psychiatric symptoms over the life course. Our approach to this issue is discussed in detail later in this chapter.

Recommendation #3

- Offer to involve family members in this process to educate them and solicit their input (consent is required).

We typically ask patients to involve family members, as long as the specific family member is clearly a positive influence and understands the process of

recovery from addiction. In Ms. R.D.'s case, her father is an active alcoholic and her mother a chronic bulimic, not to mention that they are both living on the other side of the country from our patient. In cases like these, we do not involve family members but instead embark on a discussion of the environmental factors contributing toward a patient's substance use. An ecological approach to addiction treatment underscores the importance of context and social relations in the cycle of addiction.

Ms. R.D. lives in a sorority with a reputation for partying and has a circle of friends, including a boyfriend, with whom she habitually drinks to get drunk. In other words, the network of people at the center of her life use substances. It is essential to explore with substance use–disordered patients how the relationships in their lives will be affected if they stop using substances. Not uncommonly, removing the major shared activity—and one that at least part of the time is quite rewarding—can threaten the foundation of the relationship. We try to acknowledge that stopping substances may have a destabilizing effect on the patient's social milieu. Likewise, we ask patients to identify who in their lives might help them stop using substances and encourage them to seek out settings in which substances are not used, which may range from religious organizations to cultural/civic clubs to athletic teams or study groups. Depending on the arrangements available on the campus, another possibility when necessary is to shift living arrangements to a dry floor in a dormitory.

Recommendation #4

- Assess patient's degree of readiness for change (e.g., "How willing are you to consider reducing your use at this time?").

Ms. R.D. is willing to make changes and agrees to a 4-week trial of abstinence from all substances. She is motivated by the prospect that her insomnia and mood swings might improve and by her own acknowledgment that her substance use is not where she would like it to be.

Recommendation #5

- Negotiate goals and strategies for change.

As mentioned earlier, we recommend that Ms. R.D. identify individuals and activities that might help her avoid substance use. We discuss triggers for substance use, such as sorority parties, and talk about whether she could avoid the parties altogether or, if not, develop strategies to abstain from substance use while attending the parties. We recommend getting rid of all substances and substance paraphernalia, including running water over her remaining cigarettes before disposing of them in her garbage can, and so on. Importantly, we emphasize to Ms. R.D. that whether or not she is successful in this 4-week trial of abstinence, we want to see her, we want to hear from her, we are here to care for her. We give her our contact information and carte blanche to call us with any questions or concerns. We warn her again that initially her symptoms of low mood, anxiety, and insomnia may get worse before they get better, due to early withdrawal, but that if she is able to remain abstinent or even substantially reduce her use, we believe she will begin to feel better in 2–4 weeks. If at that time, with abstinence, she continues to have psychiatric symptoms, we assure her that we will proceed with treatment for a mood disorder. We warn her about more serious withdrawal symptoms, such as delirium tremens and seizure, although we do not feel she is at risk for these. Although withdrawal risk is not extremely high in Ms. R.D.'s case, it is worth mentioning that in many cases, particularly those that are more severe, we often refer to the Clinical Institute Withdrawal Assessment for Alcohol (Sellers et al. 2008). This instrument screens medical and surgical patients for the risk of developing alcohol withdrawal syndrome and is published in the *Handbook of Psychiatric Measures* by Rush et al. (2008). It can also be found on the Internet.

Recommendation #6

- Treat with pharmacotherapy for addiction, if indicated.

We do not recommend or even discuss pharmacotherapy specifically for addiction with Ms. R.D. According to the VHA/DoD guideline, Ms. R.D. is a good candidate for naltrexone or disulfiram, particularly given her combined alcohol and cocaine dependence (see Table 26–3). Given that her alcohol

TABLE 26–3. Indications for using naltrexone and disulfiram for alcohol dependence, according to VHA/DoD guidelines

Naltrexone	Disulfiram[a]
Alcohol dependence with:	**Alcohol dependence with:**
Ability to achieve at least 3–5 days of abstinence to rule out the need for detoxification	Abstinence >24 hours and blood alcohol level=0
Drinking within the past 30 days and/or reports of craving	Combined cocaine and alcohol dependence
Most effective when the patient is engaged in addiction-focused counseling	Failure of or contraindication to naltrexone
	Previous response to disulfiram
	Patient preference
	Capacity to appreciate risks and benefits and to consent to treatment

Note. VHA/DoD=Veterans Health Administration/Department of Defense.
[a]Most effective with monitored administration (e.g., in clinic or with spouse or probation officer).

Source. Veterans Health Administration/Department of Defense 2004.

dependence is not extremely advanced, her readiness for change, and the lack of any recent prior quit attempts, we decide to bypass addiction pharmacotherapy, at least initially. However, we do offer low-dose trazodone, an antidepressant which is an effective, nonaddictive sleep aid. She declines.

Recommendation #7

- Schedule an initial follow-up appointment in 2–4 weeks.

Given the severity of Ms. R.D.'s substance use, including a severe binge drinking pattern and the use of illicit drugs as well as her passive suicidal ideation, we feel it is important to see her back within a week and to establish a safety contract for her to call us if she gets much worse in the interim. Ms. R.D. agrees to this plan, yet she does not appear for her next appointment.

We frequently experience this phenomenon: A patient expresses willingness to a trial of reduced substance use or abstinence and then does not appear at the follow-up appointment. It represents an important juncture in the treatment of addicted patients. It is essential that we call the patient at this time and emphasize to them that we want to see them, that we care about them, and that this is true even if they were not able to adhere to the contract of reduced or abstinent substance use. If the patient is not able to reduce or abstain from substances, their shame often leads them to consciously or unconsciously forget their appointment with us. Therefore it becomes essential for us to initiate contact and reemphasize our involvement in their treatment, no matter their current substance use. In our experience, a phone call like this usually brings the patient in. If the clinician does not call the patient, the patient will likely not return to treatment.

We call Ms. R.D., and she states that she had forgotten the appointment but says she is doing better, has drastically reduced her alcohol consumption, and is not using any other substances except for "occasional" nicotine. She is unable to be specific as to the precise amounts of substances consumed, although we believe her report of substantial reduction. We emphasize how important it is for her to come in and see us, and indeed the next week she comes to her appointment. Her mood seems upbeat, and she reports that she is sleeping better than she has in years, and her mood is much steadier. She has decided to tell everyone in her sorority, and her boyfriend, that she has had to stop drinking and using drugs, and to her great surprise, they are very supportive. Her boyfriend apparently comments that he likes her "better off drugs" because she "isn't so loopy." Life is looking up, and we continue the same treatment plan, encouraging her to try to further re-

duce her alcohol consumption down to nothing and to even try to abstain from cigarettes, which she is now smoking at the rate of three or four per week. We schedule an appointment for 2 weeks from now.

Recommendation #8

- Offer and recommend smoking cessation treatment to every patient who is dependent on nicotine. Use the VHA/DoD *Clinical Practice Guideline for the Management of Tobacco Use* (Veterans Health Administration/Department of Defense 2004).

Nicotine dependence is a diagnosis that is too often ignored and minimized by mental health care providers (Himelhoch and Daumit 2003). We feel strongly that nicotine dependence should be included in the diagnoses and aggressively treated, and this viewpoint is supported by the VHA/DoD guideline as well as most other guidelines. Although most clinicians would have stronger concern about alcohol and cocaine use in cases like that of Ms. R.D., the epidemiologic reality is that, on average, nicotine is far more likely to result in death over the lifespan of such individuals. Nicotine addiction is the highest contributor toward morbidity and mortality in this country, taking up to 350,000 lives every year. It is important for mental health care providers to identify nicotine dependence as a health care issue and to discuss treatment options, lest they covertly communicate that smoking is not a problem (Lembke et al. 2007).

Ms. R.D. is initially able to reduce her daily cigarette consumption but not cut cigarettes out altogether. We take time at her third visit to address nicotine again and to encourage her to stop smoking completely. We offer a U.S. Food and Drug Administration–approved smoking cessation remedy, namely bupropion (Wellbutrin or Zyban) and nicotine replacement therapy, and encourage her to set a specific quit date, both interventions that have been associated with the best outcomes. Ms. R.D. accepts a smoking cessation trial intervention with bupropion; however, 3 days into treatment with bupropion, prior to quitting smoking, she calls to report she is feeling very jittery and "amped" and is unable to sleep. Her mood is euthymic with brief period of feeling down, lasting less than half a day. She has maintained a lower level of alcohol consumption and continues to be abstinent from cocaine and marijuana. We immediately call her back and suggest she stop the bupropion until further discussion at her next visit, scheduled for the next week.

Recommendation #9

- Monitor changes at follow-up visits by asking patient about use, health effects, and barriers to change.

When Ms. R.D. presents for her fourth appointment, 5 weeks from her initial presentation, she reports that her sleep has continued to be poor, even after stopping the bupropion, although still not as bad as before she initially presented. She also reports that her mood is consistently low nearly all day, 2–3 days a week. She is not suicidal, she has continued to abstain from marijuana and cocaine, and her alcohol consumption has gone down to zero for the past week, in the hope that complete abstinence would buoy her mood. Despite these efforts, she finds herself getting progressively more depressed. She is still smoking, and even increased her cigarette use to five or six per week, or approximately one per day. Although we are of course concerned about Ms. R.D.'s downturn in mood, we are very encouraged that she called us about it and that she continues to refrain from hazardous substance use, because it suggests we have been successful in making a strong therapeutic alliance, despite a somewhat rocky start.

Recommendation #10

- Treat concurrent psychiatric disorders, including concurrent pharmacotherapy.

There is no evidence base by which to judge how long a patient needs to abstain from substances before one can determine whether psychiatric symptoms are substance induced or represent a *forme fruste* mental illness other than a substance use disorder. DSM-IV-TR recommends a minimum of 4 weeks, and in working in our clinic, we like to see a minimum of 4 weeks of abstinence (or at worst, minimal use) before we make the diagnosis. Every case is different, however, and there is no hard-and-fast rule that can eliminate the need for clinical judgment when one is making these difficult determina-

tions. In the case of Ms. R.D., we feel that her downturn in mood after 4–5 weeks of significant substance reduction and abstinence from many substances she had previously been abusing, along with her long history of mood instability, is enough to convince us that she probably has a mood disorder as well. Given her predominantly depressed mood with prominent anxiety and insomnia, and her poor reaction to bupropion, we decide to start her on the selective serotonin reuptake inhibitor citalopram.

She responds well, with improvements in sleep, mood, and anxiety. She remains largely abstinent from substances, and when she does use alcohol and/or nicotine she does so in moderation. She continues to struggle with affect regulation and self-esteem, and for these issues we refer her to psychotherapy and a skills training group, which emphasizes distress tolerance and mindfulness meditation. She does well for 8 months and then completely disappears from treatment after the summer break from school. We have not seen her in several years.

Summary

The patient presents with a chief complaint of mood instability marked most recently by depression with fleeting suicidal ideation, but no imminent threat to herself or others. A careful screening for alcohol consumption, focusing initially on amounts and frequency in the past week using the TLFB, reveals a level of alcohol consumption far above what is considered safe and well within the hazardous range. This discovery prompts an in-depth substance use screening and history, which in turn reveals nicotine dependence, cocaine dependence, marijuana abuse, and probably cocaine withdrawal. Keeping the patient's chief complaint in mind, we give feedback about the substance use screening results, educate her about safe levels of drinking, challenge normalization of her substance use, and negotiate strategies for change, using her chief complaint as incentive to target substance use. Four to five weeks after the patient commits to change and markedly reduces her substance use, she still experiences symptoms of depression, at which point we begin treatment with an antidepressant medication, even as we continue to assess and intervene in the substance use problems. We also offer a psychotropic agent for smoking cessation treatment, but the patient does not tolerate the medication and discontinues it. Four to five months into treatment, the patient's mood and functioning is significantly improved, and she has largely eliminated all substance use, except for the occasional alcoholic beverage and cigarette. She continues to do well for the entirety of the academic year but is lost to treatment after the summer break.

Ways to Improve Practice

We made a significant error in this case by not working out a system for the patient to track her consumption during treatment. We were thus left to accept her account of "substantial reduction" in alcohol consumption and "occasional" use of nicotine without knowing the exact meaning of these terms. Careful monitoring is desirable in several respects (Kanfer and Scheft 1988). First, monitoring reduces the behavior through reactance. Second, it allows both the patient and provider to see change over time, which can be motivating and informative for care planning. Patients like Ms. R.D. who drink in social situations are often averse to monitoring, and in such cases it is useful to develop a covert strategy, for example, moving a coin from the right pocket to the left for each drink that is taken.

References

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

Danielsson PE, Rivara FP, Gentilello LM, et al: Reasons why trauma surgeons fail to screen for alcohol problems. Arch Surg 134:564–568, 1999

Himelhoch S, Daumit G: To whom do psychiatrists offer smoking-cessation counseling? Am J Psychiatry 160:2228–2230, 2003

Humphreys K: Circles of Recovery: Self-Help Organizations for Addictions. Cambridge, UK, Cambridge University Press, 2004

Kanfer FH, Scheft BK: Guiding the Process of Therapeutic Change. Champaign, IL, Research Press, 1988

Kessler RC, Nelson CB, McGonagle KA, et al: The epidemiology of co-occurring addictive and mental disorders: implications for prevention and service utilization. Am J Orthopsychiatry 66:17–31, 1996

Kessler RC, Berglund P, Demler O, et al: The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 289:3095–3105, 2003

Lembke A, Johnson K, DeBattista C: Depression and smoking cessation: does the evidence support psychiatric practice? Neuropsychiatr Dis Treat 3:1–7, 2007

Mueser KT, Drake RE, Wallach MA: Dual diagnosis: a review of etiological theories. Addict Behav 23:717–734, 1998

Nunes EV, Levin FR: Treatment of depression in patients with alcohol or other drug dependence: a meta-analysis. JAMA 291:1887–1896, 2004

Rush AJ, First MB, Blacker D (eds): Handbook of Psychiatric Measures, 2nd Edition. Washington, DC, American Psychiatric Publishing, 2008

Sellers EM, Sullivan JT, Somer G, et al: Clinical Institute Withdrawal Assessment for Alcohol (CIWA-AD), in Handbook of Psychiatric Measures, 2nd Edition. Edited by Rush AJ, First MB, and Blacker D. Washington, DC, American Psychiatric Publishing, 2008, pp 460–461

Sobell LC, Sobell MB: Alcohol Timeline Followback Users' Manual. Toronto, ON, Canada, Addiction Research Foundation, 1995

Substance Abuse and Mental Health Services Administration [SAMHSA]: Results from the 2007 National Survey on Drug Use and Health: national findings (NSDUH Series H-34, DHHS Publ No SMA 08–4343). Rockville, MD, Substance Abuse and Mental Health Services Administration, Office of Applied Studies, 2008

Swift RM: Drug therapy for alcohol dependence. N Engl J Med 340:1482–1490, 1999

U.S. Preventive Services Task Force: Guide to Clinical Preventive Services, 2nd Edition. Washington, DC, Office of Disease Prevention and Health Promotion, 1996

Veterans Health Administration/Department of Defense: Clinical Practice Guideline for the Management of Substance Use Disorders. VA Office of Quality and Performance. Washington, DC, Department of Veterans Affairs/Department of Defense, 2001

Veterans Health Administration/Department of Defense: Clinical Practice Guideline for the Management of Tobacco Use. Management of Tobacco Use Working Group. Washington, DC, Department of Veterans Affairs, 2002. Available at: http://www.healthquality.va.gov/tuc/tuc_fulltext.pdf

Veterans Health Administration/Department of Defense: Clinical Practice Guideline for the Management of Substance Use Disorders: Guideline Summary, v.2.0. Management of SUD Working Group. Washington, DC, Department of Veterans Affairs, 2009. Available at: http://www.healthquality.va.gov/sud/sud_sum_v2_0.pdf

This page intentionally left blank

27

A Complex Personality Disorder Case

James Reich, M.D.

Setting

I work as a solo practitioner doing outpatient psychiatry. My specialty is difficult-to-treat anxiety disorders, which are often comorbid with personality disorders.

Illustration

This study illustrates:

- Management of a complex case of two personality disorders as well as a comorbid Axis I disorder, obsessive-compulsive disorder (OCD)
- Use of expert guidelines for borderline personality disorder (American Psychiatric Association 2001) and OCD (American Psychiatric Association 2007)
- Psychopharmacology of personality disorders (Herpertz et al. 2007)

Chief Complaint

The patient, Mr. X, presents as a 30-year-old single male with significant social avoidance, obsessive thoughts, and impulsive behavior. He wishes to stabilize his symptoms so he can pursue work.

Present Illness

Mr. X was a 30-year-old single lawyer. His obsessions began in his teenage years but did not interfere with his ability to finish high school or, ultimately, college and law school. However, he had had unstable relationships with sudden shifts between feeling very positive toward someone and feeling that the same person had no redeeming qualities. This caused difficulty at work, although he tried to be professional. He had some identity disturbance in which he questioned who he was, including having doubts about his sexuality. He had had impulsive behavior with regard to alcohol and sex. At times he would feel intense anger. At one point he faced the possibility of serious legal repercussions for assault, but fortunately this did not happen. He had intense dysphoric anxious episodes with extreme self-doubt, which would go away only to recur. He also had chronic feelings of emptiness.

Mr. X had an intense fear of rejection. It was hard for him to follow up on job leads; often he would call back when he doubted there would be someone to answer. He was intensely afraid of being criticized at job interviews, leading him to postpone or cancel them. His social insecurity affected his ability to date, and he was not dating despite opportunities to do so.

He had some strengths as well. In addition to completing his schooling, Mr. X had worked at a law firm for several years and gotten some good experience. He was capable of forming friendships, though this was difficult for him, and seemed to be liked by his friends. He had a good relationship with his family.

Mr. X had worked with a number of psychiatrists and psychologists intermittently. Medication classes he had been on included antidepressants, antipsychotics, antianxiety agents, and mood stabilizers. In

general, the medications had been given during periods of significant exacerbation of symptoms at relatively high dosages and then tapered off. At the time he presented to me he was on a very low dose of a selective serotonin reuptake inhibitor (SSRI) and a medication for sleep as needed.

Other Significant Findings From Assessment

Mr. X had had periods of depression that were treated with antidepressants, raising the possibility of major depressive disorder.

His binge drinking, although not frequent, could involve as many as 10 drinks in an evening and result in risky behavior.

His lack of structure and his anxiety had resulted in a somewhat reversed sleep-wake cycle.

Differential Diagnosis

Most of the diagnostic issues were fairly straightforward. The only real question was whether some of the personality symptoms were due to state effects that would be reduced by treatment of the comorbid Axis I disorder(s) (Reich 2007).

DSM-IV-TR Diagnosis

Axis I	OCD
	Alcohol abuse
	Rule out major depression
Axis II	Borderline personality disorder
	Avoidant personality disorder
Axis III	No significant medical problems
Axis IV	Unemployment, lack of structure
Axis V	GAF score: 52

Treatment Plan Considerations

Primary Problems and Guideline Selection

Personality pathology was clearly a major problem for this patient. There are no algorithms for the treatment of personality disorder in general, but the American Psychiatric Association does have guidelines for the treatment of borderline personality disorder (Oldham et al. 2001). In addition, the World Federation of Societies of Biological Psychiatry has guidelines for the biological treatment of personality disorders (Herpertz et al. 2007)

The obsessive symptoms were also clearly a major problem. Here we follow the American Psychiatric Association guidelines for treatment (Koran et al. 2007).

Secondary Problems and Issues

Of concern is Mr. X's episodic binge drinking. Although not frequent, it could have significant life consequences. Another secondary issue is his depression, which although not current would represent a potential future problem.

Treatment Goals, Measures, and Methods

Table 27–1 shows the treatment goals, measures, and methods for Mr. X's case.

Course

The guidelines for borderline personality disorder and OCD, cited earlier, emphasize the formation of a good therapeutic relationship with the patient, stressing psychotherapy. Mr. X and I set up weekly meetings and decided on a cognitive-behavioral approach, which fit both the treatment guidelines for borderline personality disorder and OCD and the patient's own preferences. Initially we made no changes to his medications, but we spent time understanding Mr. X's symptoms and establishing a relationship.

It became clear that Mr. X had such overwhelming anxiety that we would need to proceed with some medication relief to allow him to be calm enough to make progress. He had had sexual side effects on increased SSRI dosage in the past and was reluctant to try another antipsychotic, which he also had tried in the past. Benzodiazepines were contraindicated by the borderline personality disorder treatment guidelines because they could exacerbate impulsive behavior and might not be helpful for binge drinking. The World Federation of Societies of Biological Psychiatry guidelines (www.wfsbp.org/treatment-guidelines.html) indicated that avoidant personality disorder might be treated similarly to generalized social phobia, which would respond to benzodiazepines (Reich 2009). Ultimately we tried a combination of clonazepam 0.5 mg in the morning and alcohol education, alcohol counseling, and very careful alcohol and drug intake monitoring.

TABLE 27–1. Treatment goals, measures and methods

Treatment goal	Measure	Method
Establish a treatment relationship to work together in psychotherapy (ability to communicate difficulties and collaboratively work out plans to approach them)	Attendance and treatment compliance	Weekly cognitive and behavioral psychotherapy
Reduce personality symptoms characterized by avoidance	Ability to follow through in feared situations such as job interviews[a]	Medication and cognitive and behavioral therapy
Reduce personality symptoms characterized by impulsivity and anger	Episodes of socially inappropriate anger or impulsive behavior	Medication and developing alternate strategies to deal with uncomfortable feelings
Reduce obsessive symptoms	Level of obsessive thoughts[a]	Medication and cognitive and behavioral therapy
Reduce frequency of episodic binge drinking	Episodes of binge drinking	Alcohol education and counseling

[a]Measured on a nonstandardized 0–10 scale.

The use of a benzodiazepine in a patient with impulse-control problems and binge drinking was clearly a difficult risk management decision. I chose a longer-acting benzodiazepine to reduce addictive potential and used this medication in small regular doses. There were to be no "as needed" doses, and Mr. X understood that any escalation in dosage or change in drinking patterns would result in stopping the clonazepam. To ensure compliance, prescription amounts and dates were closely monitored. Also, at least at first, each session would involve discussion of alcohol and/or use and relevant education.

Mr. X's anxiety was reduced enough to focus on some of the OCD symptoms. Following medication treatment consistent with both borderline and OCD guidelines we increased his SSRI incrementally to 100 mg (where side effects precluded going higher) and ultimately added the antipsychotic quetiapine 50 mg at night. The quetiapine helped both his OCD symptoms and his reversed day/night cycle and was consistent with borderline personality disorder guidelines. The medication was combined with behavioral and cognitive techniques, which he was able to apply reasonably well. We used these same techniques to work on his social avoidance. I felt this was an area in which progress could be made and that would give Mr. X more self-confidence.

The alcohol treatment started with Mr. X coming to the conclusion he would want to stop drinking at some point (not then). He tried limiting his drinking in social situations but was not very successful. He made an attempt at abstinence that failed. Later, he made a second attempt at abstinence that went better and that he thought helped his overall mental health.

His functional status improved somewhat, and he signed up for education classes to bolster his areas of expertise, but progress was not as fast as we hoped. Very far into the therapy Mr. X brought up a sexual trauma that he had experienced as a teenager, which seemed to have a connection with his impulsivity and anger. The therapy then focused on this area, which is ongoing.

The goals to establish a relationship with a significant other and return to the workplace are works in progress.

Tables of Improvement

Tables 27–2 through 27–6 describe the patient's improvement during the course of treatment.

Ways to Improve Practice

There are real challenges to using guidelines when a patient has multiple significant problems. These challenges include lack of guidelines (i.e., treatment of avoidant personality disorder or a mixture of personality disorders); the possibility of conflicting advice as to the use of a particular technique (e.g., use

TABLE 27–2. Establishment of treatment relationship

Weeks 0–10	Weeks 11–20	Week 21 onward
Occasional missed appointments without agreed prior notice	Some changes in medication use without consultation (mainly discontinuation due to side effects)	Reasonably good attendance and discussion of medication and therapeutic issues

TABLE 27–3. Reduction of personality symptoms characterized by avoidance

Weeks 2–10	Weeks 11–20	Weeks 21–40	Week 41 onward
Trials and discussion of trials of different medications (many of which did not work or could not be tolerated) Score=9	Some successful adjustment of current medications (selective serotonin reuptake inhibitor) and addition of new medications (benzodiazepine and atypical antipsychotic) Score=6.5	Largely therapy to encourage use of cognitive-behavioral techniques Score=5	Use of earlier techniques plus discussion and self exposure of feelings of defectiveness related to teenage sexual trauma Score=4

TABLE 27–4. Reduction of personality symptoms characterized by impulsivity and anger

Weeks 0–10	Weeks 11–40	Week 41 onward
Review of prior anger and impulsive actions with discussion on their effects on client's life Client always recognized the destructive aspects of these incidents Score=6	Discussed and examined anger and impulse issues as they came up; seemed to be improvement with reduction of other symptoms and especially with gradual control of alcohol use Score=3	Added to earlier treatment was discussion of how early sexual trauma created feelings of defectiveness and how this related to anger and impulsive actions Score=2

TABLE 27–5. Reduction of obsessive symptoms

Weeks 0–10	Weeks 11–20	Weeks 21–30	Weeks 31–40	Week 41 onward
Trials and discussion of different medications (many of which did not work or could not be tolerated) Score=9	Some successful adjustment of current medications (selective serotonin reuptake inhibitor) and addition of new medications (atypical antipsychotic) Score=6	Cognitive and behavioral techniques with much emphasis on exposure and response prevention Score=4	Continued use of all of the earlier techniques Score=3	Specific discussions of sexual obsessions Score=2

TABLE 27–6. Reduction of frequency of episodic binge drinking

Weeks 0–5	Weeks 6–15	Weeks 16–20	Week 21 onward
Assessment of problem and education Client believed he might someday wish to stop drinking but at present wished to try to become a "social drinker"	Continued assessment of motivation to change and education Social drinking not successful; several significant binges	Client decided to try abstinence but rejected formal programs such as Alcoholics Anonymous At week 20 attended wedding with old college drinking friends—binge drinking	Therapy for fears related to not drinking (will have no social life, etc.) and specific behaviors to avoid the beginning of a binge Client still rejects formal programs Abstinent so far

of benzodiazepines for avoidant personality disorder vs. borderline personality disorder); and prioritizing which steps to take first.

Nonetheless I found the guidelines quite helpful. When all the guidelines agreed on certain interventions, these obviously rose to the top. In this case all of the guidelines emphasized the importance of forming a therapeutic relationship and the use of at least two categories of medication (SSRIs and atypical antipsychotics). Even when the guidelines conflicted to some extent, they helped in the formation of a risk-benefit decision.

The guidelines can also help us reflect on our own practice patterns. For myself, this case showed me how hard it is not to use benzodiazepines in cases with considerable anxiety even when there are contraindications.

Although the guidelines do not anticipate every contingency and do not replace clinical judgment (the emergence of sexual trauma as an issue in this case came as a surprise to me and was not really anticipated by the various guidelines), they do provide a solid base of advice and a measure to examine your own treatment approaches.

References

American Psychiatric Association: Practice guidelines for the treatment of patients with borderline personality disorder. Am J Psychiatry 158:1–52, 2001

American Psychiatric Association: Practice Guideline for the Treatment of Patients With Obsessive-Compulsive Disorder. Washington, DC, American Psychiatric Association, 2007. Available at: http://www.psychiatryonline.com/pracGuide/pracGuideChapToc_10.aspx. Accessed September 18, 2009

Herpertz SC, Zanarini M, Schultz CS, et al: World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of personality disorders. World J Biol Psychiatry 8:212–244, 2007

Koran LM, Hanna GL, Hollander E, et al: Practice guideline for the treatment of patients with obsessive-compulsive disorder. Am J Psychiatry 164(suppl):1–53, 2007. Available at: http://www.psychiatryonline.com/pracGuide/pracGuideTopic_10.aspx. Accessed June 18, 2009

Reich J: State and trait in personality disorders. Ann Clin Psychiatry 19:37–44, 2007

Reich J: Avoidant personality disorder and its relationship to social phobia. Curr Psychiatry Rep 11:89–93, 2009

This page intentionally left blank

28

Bulimia Nervosa

Joanna M. Marino, M.A.
James E. Mitchell, M.D.

Setting

Patients with bulimia nervosa, most of whom present for treatment in their late teens or 20s, generally can be treated successfully in outpatient settings. Exceptions include patients who are medically unstable; patients for whom comorbid conditions, such as severe depression and suicidality, dictate the need for inpatient treatment; and patients with severe comorbid personality problems who may benefit from a structured partial hospitalization program. However, for the average patient, outpatient treatment either in an individual or group format is both adequate and preferable. Our treatment setting includes a team approach composed of psychologists, dietitians, a physician, and advanced practice nurses, because this approach targets the psychological, behavioral, nutritional, and medical needs of eating disorder patients.

Illustration

The following case provides an example for discussion of the application of the American Psychiatric Association (2006) *Practice Guideline for the Treatment of Patients With Eating Disorders*. Initial assessment of eating disorder symptoms should include height, weight, and body mass index (BMI) measurements. The assessment should review the patient's history of eating behaviors, compensatory behaviors, and beliefs about weight, shape, and food. Gathering a general psychosocial history in order to identify stressors and assess comorbid psychopathology provides helpful information for the course and outcomes of treatment.

Chief Complaint

Allison is a 22-year-old female who presents with complaints of depressed mood, anxiety, and binge eating and purging episodes.

Present Illness

Allison reported her binge eating and vomiting developed at age 16 and seemed to help her "relax," especially in the evenings when her binge/purge episodes typically occurred. She reported binge eating and vomiting up to four times each evening. She hoped to avoid these episodes but was afraid of gaining weight. Her current weight was 120 lb and her height was 5'4" (BMI=20.6).

Past Psychiatric History

The patient reported a history of depressed mood and anxiety since age 15. She denied a specific event that precipitated her mood symptoms. She had had suicidal ideation at age 15, although she reported no current plan or intent. At age 15, she often felt as if she was "not good enough or pretty enough," especially when compared with other girls at school. She experienced bouts of unprovoked crying, anhedonia, concentration difficulties, and worry about the future, which caused her to have difficulty falling asleep. She reported that these symptoms have continued since

age 15 but "aren't as intense" as they were when they first occurred. She denied past psychological treatment for any mental health concerns.

At age 15, Allison began to gain weight. Her highest weight was 140 lb. She then began to diet, reducing her intake to 1,200 calories per day. She reported that this dieting lasted about a month and that she lost approximately 8 lb. On a Saturday when her parents were away, she experienced a binge eating episode. She then developed a pattern of binge eating every weekend and then restarting her dieting pattern every Monday. After about 2 months in this pattern, Allison experienced a binge eating episode that was larger than her prior episodes. She felt the urge to vomit in order to "feel just a little less stuffed." Vomiting came easily to Allison and she thought she could be successful at losing more weight if she vomited after each dinner. Allison's binge/purge cycle then became more frequent and was occurring at least once per day by the time she was 17.

History of Substance Abuse

The patient reported drinking one or two alcoholic drinks once monthly. She smoked marijuana on three occasions in tenth grade and related this to "peer pressure." She denied use of any other substances.

Exercise, Diet, and Stress Management

Allison exercised five times weekly for 1.5 hours during each episode. Her exercise routine included running for 60 minutes and a weight-lifting routine lasting 30 minutes. She reported exercising even though on most days she felt fatigued. She had had some dizziness when running. Her meal plan was also quite stringent, including a diet of 1,500 calories per day and limited fat intake. However, on most days she reported being unable to follow her diet and feeling "ravenous" when arriving home from work at 6 P.M. She did not experience the urge to binge eat during the day; however, when leaving work she became anxious about her performance during the day. She reported intrusive thoughts of binge eating on her drive home from work and feeling compelled to stop at the grocery store for a few items to prepare for dinner. Her binge eating episodes always involved high-fat food items such as pizza, cookies, chocolate milk, and chips.

Allison reported eating quickly and feeling as if she "blacked out" when she had a binge eating episode. She stated, "I know what I'm doing, bingeing on all that food, I just can't stop." She became overwhelmed with a sense of guilt after her eating episodes and then vomited several times. The vomiting decreased her anxiety and "clears my head… I don't have to worry when I'm binge eating and I can then just get rid of the food." Allison related her depressed mood to her inability to control her food intake and "feeling fat." She desired to weigh 110 lb (BMI=18.9).

Social History

Allison reported three or four close friends throughout her schooling and several acquaintances. She had been involved in soccer and tennis during eighth grade; however, she quit both teams to pursue her interest in playing piano, drawing, and painting. At the time of her evaluation she had very few friends. She denied involvement in a romantic relationship.

Educational and Occupational History

The patient completed the twelfth grade and attained her associate's degree in graphic design at a local community college. She worked at a fast food restaurant during high school and then began working at a well-known local graphic design studio after meeting the owner through an acquaintance.

Family History

Allison was raised by her biological mother and father. She had one older brother who was studying medicine. She reported that her upbringing was "good…just typical, I guess." Her father was employed as an executive at a local grocery store, and her mother was an administrative assistant. She reported a "good" relationship with her brother and parents but suggested her mother was sometimes critical of her, accepting only "A" grades as "good." Allison was on the "B honor roll," and this was often a point of contention with her mother. Her mother was often concerned about what other members of her church thought about the family and was focused on appearing "worry-free…or perfect."

Other Significant Findings From Assessment

Review of Systems

- Unremarkable except for occasional headaches, primarily frontal and usually relieved by aspirin
- Occasional palpitations and tachycardia post vomiting
- Problems with intermittent constipation and diarrhea; occasional upper abdominal pain associated with binge eating
- Had two episodes in which she found trace amounts of blood in her vomitus

Physical Examination

Patient was a well-developed, well-nourished female in no acute distress. Her blood pressure was 110 over 62, right arm sitting; her pulse was 64 and regular; and her respiratory rate was 12. Physical examination was essentially negative except for some evidence of scar formation on the dorsum of her right hand where she had traumatized the skin while self-inducing vomiting. This is known as Russell's sign.

Diagnostic Tests

Allison's diagnostic tests were within normal limits. Of particular importance in this testing is obtaining a serum electrolyte determination, because this is the blood chemistry most commonly affected by bulimia nervosa. Of particular concern is the risk of severe hypochloremia, metabolic alkalosis, and hypocholemia.

Psychological Tests

Allison completed the Beck Depression Inventory–II and attained a raw score of 21, suggesting moderate depressive symptoms. She also completed the Beck Anxiety Inventory with a score of 30, suggesting moderate symptoms of anxiety. The results of Allison's Eating Disorder Inventory–3 revealed clinical elevations (i.e., T/Composite Score≥50) on Drive for Thinness (T=52), Bulimia (T=56), Interpersonal Alienation (T=61), Personal Alienation (T=60), Perfectionism (T=58), and Affective Problems (T=58). The elevations appeared to align with Allison's reported dieting and binge/purge behavior as well as her lack of social support, perfectionistic tendencies, and mood problems.

DSM-IV-TR Diagnosis

Axis I	Bulimia nervosa Major depressive disorder, single episode, moderate Anxiety disorder not otherwise specified
Axis II	No diagnosis on Axis II
Axis III	Patient reports headaches; see medical record
Axis IV	Occupational stressors, limited social support
Axis V	GAF score: 60

Treatment Plan Considerations

In general, treatments for bulimia nervosa have focused on pharmacological and psychotherapeutic approaches. Antidepressants were first used for the treatment of this condition because of the observation that many patients with bulimia were comorbidly depressed, and it was assumed that if their depression improved their eating disorder would improve as well. However, research has shown that the presence of depression does not predict response to antidepressant treatment in these patients. A variety of antidepressants have been studied, including monoamine oxidase inhibitors, tricyclic antidepressants, and most recently serotonin reuptake inhibitors. Currently the only drug approved by the U.S. Food and Drug Administration for the treatment of bulimia nervosa is fluoxetine. It is notable that the drug seems to work best in high dosages, at approximately 60 mg/day. Many patients will tolerate this as the initial dosage. By analogy, most practitioners prescribe fairly high dosages of other antidepressants if alternative medications are being used. The controlled treatment literature suggests that most of these drugs work reasonably well. However, fluvoxamine may be ineffective, and bupropion should be avoided because it seems to have a high propensity for causing seizures in this patient population.

Relative to psychotherapeutic interventions, a variety of psychotherapies have been described in the literature, but much of the work has focused on the use of cognitive-behavioral techniques in either group or individual formats. It is safe to conclude that currently cognitive-behavioral therapy (CBT) is the treatment of choice for bulimia nervosa. This treatment is usually delivered in a twice-weekly format for the first month or so, and this seems to be an

important variable in determining treatment outcome. There is also literature to suggest that interpersonal therapy may be effective, although the data are more limited and the treatment response seems to be somewhat delayed. In addition, there is some indication that dialectical behavior therapy can be helpful for some patients.

The question often arises as to whether antidepressant therapy and CBT should routinely be used in combination in the initial treatment of patients. Studies that have examined this question have found some modest benefit for the addition of antidepressants to CBT, but it is not clear that this outweighs the added costs and risks involved. Many practitioners recommend beginning with CBT and adding antidepressant treatment if there is no evidence of a fairly prompt response, demonstrated by reductions in the frequency of targeted behaviors, early in treatment.

As mentioned, individual and group psychotherapeutic interventions have been helpful in treating patients with bulimia nervosa. Dietary consultation is also beneficial in combination with psychotherapy. Family-based interventions are especially useful for adolescents because interpersonal issues can be discussed in the context of eating-disordered behavior. Self-help and support groups are also being studied for their effectiveness.

The American Psychiatric Association guideline's suggested aims for the treatment of individuals with bulimia nervosa are incorporated into Table 28–1. Note that the treatment interventions utilized target pharmacological, medical, and psychological interventions. A team approach to treating patients with bulimia is valuable, especially when considering the need for continued medication management and psychological interventions. Additionally, patients with medical or psychology comorbidities will need continued assessment to determine the appropriate level of care, especially if the patient fails to respond to outpatient treatment. The treatment guideline provides detailed information regarding the appropriate level of care for eating disorder patients based on several factors, including age, control over eating disorder behaviors, medical status, and location.

The outpatient clinic where Allison received her psychological treatment was composed of psychologists and psychiatrists. After the initial assessment, the importance of a treatment team was discussed with Allison. She was given referrals to a dietitian and a primary physician. The team of providers had worked together for several patients and had a strong history of good communication about patient needs and progress. This open communication was an essential component, especially when she became resistant to normalizing her eating out of fear of gaining weight. The dietitian communicated with the therapist through a hospital-wide computer system about progress and topics to address in therapy. Allison was referred to a dentist outside of the hospital system to address dental erosion secondary to her purging behavior.

Allison's initial assessment revealed comorbid anxiety and depression, which is common in eating disorder patients. Medication use can be implemented at the beginning of treatment or as treatment progresses, depending on the severity of the comorbid psychological condition. It is important to consider that mood and anxiety symptoms, especially depression, can remit as binge/purge episodes subside, providing one reason for delaying the use of pharmacological agents.

In addition to her psychological assessment, Allison was seen by a primary practice physician. Results of her medical assessment were largely within normal limits. Given her medical stability, Allison was a good candidate for outpatient treatment. She was followed monthly by the primary physician to ensure medical stability.

Treatment Goals, Measures, and Methods

The goals, measures, and methods for this treatment are outlined in Table 28–1.

Course

Allison met with the therapist twice weekly for the first 4 weeks of therapy, as is often indicated in CBT for bulimia nervosa. The beginning of psychotherapy was marked by a focus on developing a therapeutic alliance. The therapist validated the role the eating disorder behaviors served for Allison and provided a supportive and open environment to discuss concerns about the treatment course. This allowed for an environment in which Allison would feel comfortable disclosing her eating and weight concerns. The therapist provided Allison with psychoeducation regarding the hypothesized model that

TABLE 28–1. Treatment goals, measures, and methods for a patient with bulimia nervosa

Treatment goal	Measure	Method
Reduce binge/purge episodes to zero times per week	Self-report of weekly binge/purge episodes and review of self-monitoring logs	Psychoeducation CBT Add SSRI if lack of early response
Reduce body image concerns within six sessions	Self-report, EDI-3	CBT
Reduce depression and anxiety by 50% or more by week 8	BDI-II, BAI	Consider pharmacotherapy CBT
Provide psychoeducation regarding healthy nutrition and eating patterns	Self-report	Psychoeducation Consultation with dietitian
Diminish food restriction behaviors and thoughts (e.g., "safe foods" or "dieting")	Food monitoring logs	CBT Consultation with dietitian
Increase motivation to fulfill treatment requirements	Attend all therapy sessions. Complete Motivational Enhancement Decisional Balance Sheet	Psychoeducation Treatment contract Therapeutic alliance
Treat all physical complications	Laboratory reports	Consultation with general medical provider
Discuss relapse prevention	Create a relapse prevention plan	CBT
Utilize social support by enlisting one family member or friend to support treatment	Discuss family therapy options, if applicable. Participation by family or friends during therapy sessions	Psychoeducation CBT IPT

Note. BAI=Beck Anxiety Inventory; BDI-II=Beck Depression Inventory–II; CBT=cognitive-behavioral therapy; EDI-3=Eating Disorder Inventory-3; IPT=interpersonal therapy; SSRI=selective serotonin reuptake inhibitor.
Source. Goals adapted from American Psychiatric Association 2006.

diagrammed the development and continuation of her eating disorder as a means of helping Allison understand the impact emotions, cognitions, social settings, and behaviors can have on her eating disorder. The therapist also worked with Allison to create a treatment plan, which was a way to help her feel in control of her treatment and think of herself as a member of a collaborative relationship. The practitioner also suggested to Allison that use of a self-help guide could be a valuable aid understanding eating disorders. Notably, the American Psychiatric Association treatment guideline provides a recommended list of readings that are beneficial to specific patient groups and for families of those with eating disorders. A practitioner may also consult the treatment guideline for a list of recommended reading about the CBT treatment manuals and guides.

Allison was provided with monitoring logs to detail her eating episodes throughout the day (see Figure 28–1). These monitoring logs provided valuable information about the pattern of eating Allison experienced during the day. In addition to the monitoring logs, psychoeducation regarding the importance of normalized eating was incorporated throughout the sessions. Allison was quite resistant to accepting the idea that eating throughout the day would not lead her to become "fat," and she continued to have episodes of morning restricting throughout the first half of treatment. Targeted behavioral interventions such as preplanning and packing meals was somewhat helpful in addressing her restricting behavior, although her beliefs that feeling full meant she was "fat" likely hindered her ability to change her eating patterns. Figure 28–2 depicts the reduction in Allison's binge/purge episodes. As shown, her binge/purge episodes decreased steadily, but she continued to have periodic episodes. We began discussing alternative behaviors and coping skills that Allison was able to implement when she had urges to binge eat and purge. Allison, unlike

Date: ___/___/___

Time	Intake (food/fluid)	Setting	Binge eat Y/N	Vomit Y/N	Laxative Y/N	Comments

FIGURE 28–1. Food log.

Source. Adapted from Fairburn 2008.

some patients, enjoyed monitoring her eating behavior. Some patients feel that monitoring eating behaviors makes them focus too much on their consumption habits. This tends to resolve in many patients when they begin to see the benefit of analyzing their eating behavior during the session.

In the case of Allison, her binge/purge episodes persisted and her mood remained depressed during the first 4 weeks of therapy. A selective serotonin reuptake inhibitor (SSRI) was then added at week 4, and a steady decrease in symptoms resulted. Various pharmacological interventions have been effective in treating bulimia symptoms, and an SSRI was chosen because of its general tolerability by many patients. Allison reported decreases in her depression and overall level of anxiety; however, she continued to experience anxiety related to her strenuous work environment. Relaxation techniques and coping skills were used to target these anxiety symptoms.

Allison was referred to a dietitian to address her concerns regarding "safe" foods and to gain additional information about nutrition. Allison continued to have "bad" or "unsafe" foods in her diet, which at times would trigger a binge episode. She was comfortable avoiding these foods during therapy. As the eating patterns began to stabilize over the second half of the treatment course, Allison was able to expose herself to these feared foods and then utilize her distraction behavior until the foods became more tolerable.

The final steps of Allison's treatment involved discussing a relapse prevention plan. Allison and the therapist worked on identifying triggers to a lapse in eating-disordered behavior and the use of problem-

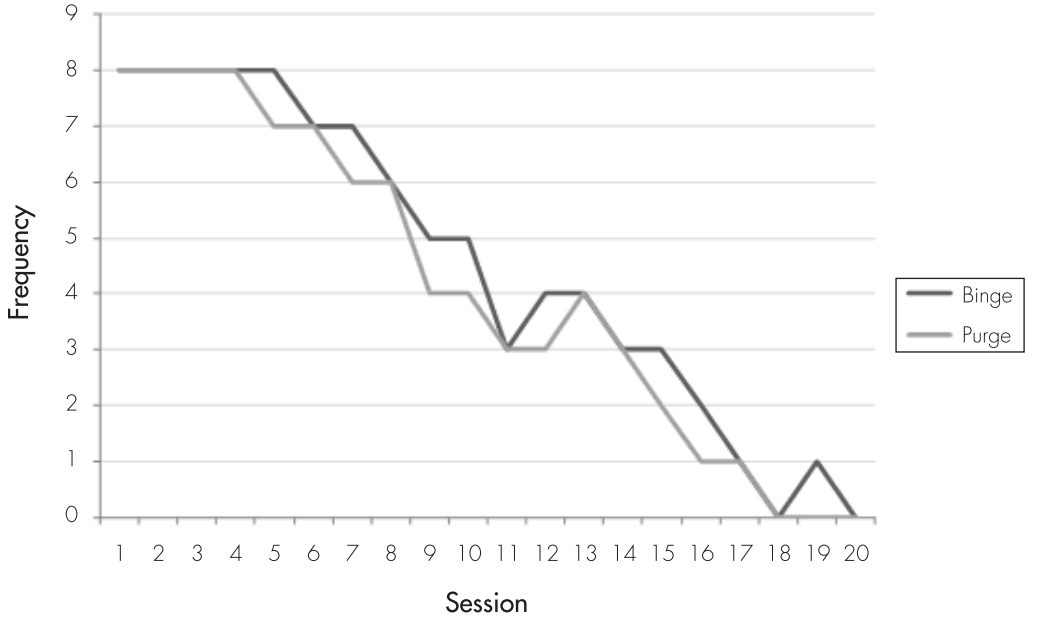


FIGURE 28–2. Binge and purge frequency over the treatment.

Sessions were biweekly during the first 4 weeks of therapy.

solving skills to avoid a full relapse. Allison became aware that minor setbacks were likely. Figure 28–2 displays the continued trouble Allison had in fully stopping her binge eating episodes. The therapeutic relationship and collaboration that was established in the beginning of treatment remained helpful in motivating Allison to continue to journey toward remission of binge/purge episodes.

In treating adolescents with eating disorders, family therapy is an important component because of the continuous interaction with parents and food in the home. In adults, involving relationship partners in adjunct therapy sessions can be beneficial, especially to aid in the monitoring of behavior or supporting behavioral changes. Treatment manuals are available for family-based therapy, and recommended readings are listed in the treatment guideline.

Allison was uninterested in involving her parents in the therapy process, even with continued discussion about the benefits of additional family therapy. Allison was very hesitant about disclosing her behavior to her family. Her situation did, however, present an ideal opportunity for the involvement of parents, especially because Allison has limited social support. She recognized that some of her anxiety and beliefs about weight were likely related to her mother's distorted beliefs about appearance and perfection. If Allison had been willing to involve her family in therapy, supplemental sessions would have occurred to address the role of families in eating disorder etiology and treatment.

Summary/Conclusion

The challenges in treating eating disorder patients are the necessities of a multidisciplinary team approach, continued assessment of medical symptomatology, and the risk of chronicity of the disorders. The treatment guideline provides valuable, explicit information regarding treatment settings and level of care (i.e., inpatient versus outpatient), medical assessments, and targets of psychotherapy. Unfortunately, not all patients are motivated for treatment, and attrition in therapy can be high. Additionally, the preponderance of eating disorder cases seem to fall into the Eating Disorder Not Otherwise Specified category, highlighting the importance of developing a treatment plan tailored to the needs of the patient.

Ways to Improve Practice

In the case described, two improvements to treatment could be considered. Allison's frequency of binge eating and purging decreased more slowly when compared with most reports in the literature regarding response patterns of bulimia nervosa patients being treated with CBT. The use of antidepressants was indicated here and was implemented in Allison's case. The addition of a third therapy session per week could also have been very helpful in reducing the frequency in these behaviors. A relapse prevention plan that included follow-up visits would also have been beneficial. Literature has suggested that patients with bulimia may not necessarily pursue follow-up treatment if they experience relapses after a successful CBT intervention (Mitchell et al. 2004). Developing a schedule of monthly follow-up sessions could be helpful in preventing full-blown relapse.

References

American Psychiatric Association: Practice Guideline for the Treatment of Patients With Eating Disorders, 3rd Edition. Washington, DC, American Psychiatric Association, 2006. Available at: http://www.psychiatryonline.com/pracGuide/pracGuideTopic_12.aspx. Accessed June 18, 2009

Beck AT, Steer RA, Brown GK: Beck Depression Inventory-II Manual. San Antonio, TX, Psychological Corporation, 1996

Fairburn CG: Cognitive Behavior Therapy and Eating Disorders. New York, Guilford, 2008

Garner DM: Eating Disorder Inventory-3 (EDI-3) Professional Manual. Lutz, FL, Psychological Assessment Resources, 2004

Mitchell JE, Agras WS, Wilson GT, et al: A trial of a relapse prevention strategy in women with bulimia nervosa who respond to cognitive-behavior therapy. Int J Eat Disord 35:549–555, 2004

Suggested Readings

Fairburn CG: Overcoming Binge Eating. New York, Guilford, 1995

Mitchell JE, Wonderlich S, Agras S: The treatment of bulimia nervosa: where are we and where are we going? Int J Eat Disord 40:94–101, 2007

Mitchell JE, Peterson CB: Assessment of Eating Disorders. New York, Guilford, 2007

National Institute for Clinical Excellence: Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders: Clinical Guideline 9. London, National Institute for Clinical Excellence, 2004

This page intentionally left blank

III

Case Studies: Evidence-Based Psychiatric Practice in Different Settings

This page intentionally left blank

29

Dual Diagnosis (PTSD/SUD) Treated in a Veterans Affairs Health Care Facility

Gretchen Lindner, Ph.D.
Steven E. Lindley, M.D., Ph.D.

Setting

Treatment was performed in a specialized outpatient posttraumatic stress disorder (PTSD) program at a Veterans Affairs (VA) health care facility. The PTSD program utilizes a multidisciplinary model wherein psychiatrists, psychologists, social workers, and nurses are all represented on the team. Veterans are typically referred to the program by other outpatient mental health programs in the hospital, psychiatric inpatient units, or primary care. The start of treatment typically involves a period of thorough assessment and treatment planning, followed by the provision of an appropriate PTSD-specific treatment. Treatment options include short-term individual therapy, short-term and long-term group therapy, couples/family therapy, medication management, and case management. The primary author was a psychology fellow in this setting and, as such, had the opportunity to work with this veteran for up to a year, functioning as the primary individual and group psychotherapist. The second author was not a treating provider but reviewed the treatment plan and outcome.

Illustration

This study illustrates:

- Application of guidelines to PTSD
- Adjustment of guidelines for comorbid PTSD and substance abuse
- Use of manualized treatments (e.g., stress management, exposure therapy, Seeking Safety)
- Concurrent use of individual therapy, group therapy, and medication management
- Coordination of care

Chief Complaint

Mr. B.H. is a 49-year-old divorced Caucasian male Army veteran who was referred to the PTSD program after he screened positive for both military sexual trauma (MST) and PTSD during a routine mental health screening in primary care. He presents to mental health with a 25-year history of intrusive memories, nightmares, insomnia, feelings of shame, and excessive alcohol use after being sexually assaulted while in the military. At the time of the rape, the veteran was 24 years old, in the U.S. Army, and stationed in Italy. While walking alone through a public park one evening, Mr. B.H. was sexually assaulted and severely beaten by five civilian men who then left him alone "to die" in the park. In addition to this sexual assault as an adult, the patient has a history of being sexually molested, between the ages of 8 and 10, by an uncle. He never discussed either the childhood molestation or military sexual assault with anyone.

Mr. B.H. states that after being assaulted in the military, he began to drink heavily in order to cope with his intense psychological reaction to the trauma. The veteran is living in a residential treatment program for homeless veterans with substance abuse problems. Although he has been clean and sober for 6 months, he continues to struggle with PTSD symptoms. He has PTSD reexperiencing symptoms that include nightmares about his assault, daily intrusive thoughts and images about the rape and his earlier molestation, and significant physical and emotional reactions in response to trauma-related cues (e.g., psychotherapy groups with other men in the program, news stories about rape, having to share a room with a man). He also has avoidance/numbing symptoms including significant avoidance of trauma-related activities (e.g., talking about his rape, touch/sex, close relationships with men), as well as numerous obsessive "safety" behaviors (e.g., showering five times a day, cleaning sheets multiple times a week). He also describes emotional numbing and severe social isolation from both friends and family. In the hyperarousal domain of PTSD symptoms, Mr. B.H. reports that he struggles with significant sleep difficulties (e.g., sleeping 2–3 hours per night), hypervigilance, irritability, and an exaggerated startle response. His score on the PTSD Checklist—Civilian Version (PCL-C) is 70.

Substance Abuse History

Mr. B.H. states that he began drinking alcohol at the age of 15, but heavy and problematic drinking began the night of his MST. He has drunk alcohol consistently for the last 25 years. Although he has had periods of productive employment (e.g., worked as a dental assistant, received numerous accolades as a program manager), the alcohol abuse progressively worsened and negatively affected his consistency at work and in relationships. He has had multiple failed marriages, long periods of unemployment, chronic homelessness, and multiple DUIs. Over the last 10 years, Mr. B.H. has enrolled in three different substance abuse programs, none of which led to substantial periods of sobriety (longest prior sobriety was 4 months). He has been participating in a homeless veterans' residential treatment program for the last 4 months and has been clean and sober for the last 6 months. The residential treatment program provides him with numerous weekly recovery-oriented group psychotherapies and weekly contact with a case manager but provides no PTSD-specific treatment.

Other Relevant Findings

Mr. B.H. reports two suicidal gestures in his life. He once played Russian roulette with a loaded gun and once lay in a bathtub with a broken bottle while strongly considering cutting his wrists.

Mr. B.H. is the youngest of four children, with two brothers and one sister. For the last 20 years he has been disconnected from his family until he recently reconnected with his sister and her children. He denies having any friends or a support system with which he feels connected, finding it difficult to trust people. He has no significant medical problems, and his medical exams have all been normal.

DSM-IV-TR Diagnosis

Axis I	Posttraumatic stress disorder Alcohol dependence, early full remission
Axis II	None
Axis III	None
Axis IV	Social isolation, unemployment, recent homelessness
Axis V	GAF score: 40

Treatment Plan Considerations

Main Problem

Mr. B.H.'s presenting problem was PTSD. Although present throughout his adult life, his symptoms appeared to have worsened significantly over his 6-month period of sobriety. He described current problems with all three clusters of PTSD symptoms (reexperiencing, avoidance/numbing, and hyperarousal).

Secondary Problem

Although the veteran was not currently using alcohol or other drugs, his long-standing history of severe dependence and his brief period of sobriety complicated treatment planning. He described a cycle of drinking in response to worsening PTSD symptoms as well as worsening PTSD symptoms during periods of sobriety. As such, there appeared to be a reciprocal relationship between his experiencing PTSD symptoms and alcohol consumption.

It was important to consider how to help the veteran develop additional coping skills to decrease his reliance on alcohol as the primary means of coping with PTSD symptoms. In addition, it was essential to consider how to provide relapse prevention work while also addressing his PTSD symptoms. Lastly, considerations of how to adjust traditional PTSD treatments to increase the likelihood of sustained sobriety were critical.

Selecting a Guideline

I (G.L.) began by reviewing the available guidelines for the treatment of PTSD, including the American Psychiatric Association (2004) practice guidelines for the treatment of patients with acute stress disorder and posttraumatic stress disorder, the Expert Consensus Guideline for treatment of posttraumatic stress disorder (Foa et al. 1999), *Effective Treatments for PTSD: Practice Guidelines From the International Society for Traumatic Stress Studies* (Foa et al. 2000), and the VA/DoD *Clinical Practice Guideline for the Management of Post-Traumatic Stress* (VA/DoD Management of Post-Traumatic Stress Working Group 2004). My predominant impression from reviewing these multiple sources was that the majority of the recommendations are similar across treatment guidelines and relied on the same body of literature. I decided to use the VA/DoD guideline in my treatment of this veteran because this guideline seemed the most appropriate given that I was working in a VA setting. This guideline also encapsulated many of the treatment recommendations highlighted in the other guideline. (See Table 29–1 for a summary of the VA/DoD guidelines and the ways in which I adhered to or made changes in the recommendations.)

The VA/DoD guideline strongly recommends several psychotherapy treatments that have been shown to have significant benefit for PTSD symptoms, including cognitive therapy, exposure therapy, stress inoculation training, and eye movement desensitization and reprocessing (EMDR). In addition, the guideline recommends using medication management as a first-line treatment for PTSD, strongly recommending the use of a selective serotonin reuptake inhibitor (SSRI), which has been shown to be effective for reducing symptom severity for all three PTSD symptom clusters, although less so in VA populations. The guideline also encourages practitioners to consider incorporating patient education and group therapy into interventions. Given these recommendations, I decided immediately to refer Mr. B.H. for medication management with one of the psychiatrists on the PTSD team. In addition, I decided on a plan for stress inoculation, exposure therapy, and cognitive therapy with this veteran. I decided against providing EMDR because the research suggests that exposure therapy and EMDR have similar outcomes, and dismantling studies suggest that the active ingredient in EMDR is the exposure (Davidson and Parker 2001). Furthermore, I have had extensive training in exposure therapy, and my clinical experience is that this is an extremely effective treatment for PTSD.

Although these seemed like appropriate treatment considerations for someone with PTSD, I also had questions about how to ensure that I was addressing treatment of his substance abuse in my plan. The VA/DoD guideline does not provide recommendations about what additional treatments, if any, should be provided when treating a veteran with comorbid substance abuse problems. As such, I went back to the research literature to identify any studies specifically addressing how to treat individuals with comorbid PTSD and substance abuse. Although traditionally the model for treatment of this population has been to first treat the substance abuse problems and then treat the PTSD problems (Brady et al. 2004), researchers have shown the effectiveness of a new manualized group-based treatment called *Seeking Safety* (Najavits 2002) that addresses both substance abuse and PTSD issues simultaneously. This 25-session group-based protocol is present centered and builds additional coping skills through cognitive, behavioral, and interpersonal modules. It is explicitly designed as a first-phase treatment of PTSD and substance dependence with a goal of increasing stabilization and safety. Given the research support for this treatment, I thought it would be appropriate to add this to my treatment plan.

Another issue during treatment planning was whether the PTSD interventions needed to be adjusted in order to treat safely those with comorbid substance abuse problems and minimize the chances of relapse. Although it is clear that exposure therapy is one of the most efficacious treatments for PTSD (Institute of Medicine 2007), the guideline gives no advice about the possible risks for those with substance abuse problems. When I was talking with treatment team members about providing exposure

TABLE 29–1. Summary of VA/DoD clinical practice guideline for the management of posttraumatic stress disorder (PTSD)

Area	Description	Followed
Interventions with significant benefit		
Medication management[a]	Medications with significant or some benefit include SSRIs, TCAs, MAOIs, sympatholytics, and novel antidepressants.	√ Referred to team psychiatrist to address medication needs
Cognitive therapy	Structured, short-term, present-centered therapy focusing on thoughts. Include thought records, thought patterns, challenging thoughts, identifying/challenging trauma-related thoughts, and distorted core beliefs.	√
Exposure therapy	Typically 8- to 12-session therapy in which patients face trauma-related feared thoughts, images, and situations through graded, repeated exposure. Can use imaginal or in-vivo exposure techniques.	√ Adjusted approach to address substance dependence history
Stress inoculation training	Teaching patients a set of skills or "toolbox" for managing anxiety and stress. The education and training of coping skills.	√
Eye movement desensitization and reprocessing	Exposure to traumatic images while shifting eyes in back-and-forth motion while also thinking of positive self-referring cognition.	Not necessary given positive response to exposure
Interventions with some benefit		
Imagery rehearsal therapy	Changes content of the patients' nightmares through imagery control.	Addressed through medications and exposure
Psychodynamic therapy	Explores psychological meaning of posttraumatic response by shifting and sorting through fantasies, fears, and defenses stirred up by traumatic event. Typically long term.	Not appropriate given VA's more short-term orientation to treatment
Other interventions to consider		
Patient education	Educate patients about the symptoms of PTSD and the multiple treatment options.	√
Group therapy	Group therapy that addresses PTSD symptoms. No clear differences/recommendations between supportive, psychodynamic, or cognitive-behavioral groups.	√

Note. MAOIs=monoamine oxidase inhibitors; TCAs=tricyclic antidepressants; SSRIs=selective serotonin reuptake inhibitors; VA/DoD=Veterans Affairs/Department of Defense.

[a]There are additional explicit guidelines for medication management, not summarized here because the primary author referred to the team psychiatrist to manage psychiatric medication.

therapy with this veteran, many people voiced concerns about the high probability of relapse upon engagement in imaginal exposure. I again turned to the literature to see what the research suggests about whether the risks for relapse are too high with exposure therapy. There generally was a paucity of research in this area; however, four non–randomized controlled trial studies showed support for the benefit of exposure therapy for PTSD patients with substance dependence (Brady et al. 2001; Donovan et al. 2001; Najavits et al. 2005; Triffleman et al. 1999). A clinical observation article by Najavits and colleagues (2005) summarizes the positive results of combining Seeking Safety and exposure therapy with dual diagnosis patients. In this article they provide explicit recommendations for how to adjust exposure therapy techniques when working with someone who has a history of substance abuse. The authors suggest creating an explicit, extensive, written set of safety parameters, as well as allowing patients to process more fluidly multiple events. They also suggest encouraging patients to process both trauma and substance use memories. Lastly, the authors indicate that it is important to conduct the exposure more flexibility and to decide collaboratively on pacing and quantity of exposure. The summary of clinical experiences and research articles encouraged me to employ exposure therapy with this veteran, but with significant modifications as delineated previously.

The last major issue that arose around treatment planning was identifying the order in which the multiple treatments should be delivered. Again, the VA/DoD guidelines do not provide explicit instructions. Thus, I used my clinical intuition and experience to help me answer this question. Because Mr. B.H. was early in sobriety, I decided that it would be most appropriate to start with treatments that were present centered and skill based. In addition, the veteran indicated a preference for starting with skill building and then transitioning to trauma-focused work if the first-line interventions were not effective. Therefore we began with individual therapy that addressed both his PTSD and substance abuse issues, utilizing cognitive therapy, stress inoculation/skill building, and relapse prevention interventions. In addition, I referred him to a 12-week stress management for PTSD group that I was co-leading based on stress inoculation and cognitive-behavioral models. I would have preferred to send him directly to a Seeking Safety group for skill building, but there were no available group openings at the time of treatment planning. We therefore planned to enroll him in Seeking Safety as soon as there was an opening. Depending on his response to these interventions, I also considered exposure therapy as a follow-up treatment if his symptoms persisted despite skill building and cognitive interventions.

Treatment Goals, Measures, and Methods

Table 29–2 summarizes our treatment goals, measures, and methods for this patient.

Course

The course of our therapy is best described as having occurred in two phases. The first phase of treatment consisted of Mr. B.H. seeing me for individual therapy focused on skill building, relapse prevention, and cognitive therapy. The veteran also concurrently participated in a 12-week stress management group co-led by me and a psychology trainee. In this phase of treatment I also referred the veteran for medication management with a PTSD team psychiatrist and coordinated care with him throughout treatment. In the second phase of treatment, when the veteran's PTSD symptoms remained elevated, the veteran participated in individual exposure therapy with me as well as a Seeking Safety group co-led by myself and another psychology fellow.

Phase 1: Individual Psychotherapy and Stress Management Group

Individual Skill-Based Psychotherapy

The first step in treatment was beginning an individual, skill-based, weekly psychotherapy. Mr. B.H. and I met together 22 times before we decided to begin exposure therapy. The first two sessions consisted of PTSD assessment and treatment planning. After our second session, I also spoke to the veteran's case manager from the residential treatment program to ensure coordination of care. She indicated that she was providing him unstructured, supportive meetings, and as such his participation in a structured, PTSD-specific treatment with me would not represent a conflict or overlapping treatment.

During the subsequent 19 sessions, I utilized cognitive-behavioral interventions to help the veteran build additional relapse prevention skills (e.g., identifying triggers, developing coping plans, encour-

TABLE 29–2. Summary of treatment goals, measures, and methods

Treatment goals	Measure	Method
Increase quality of sleep by increasing number of hours of sleep per night (>5 per night) within 3 months	Self-report	Coordinating care with psychiatrist
		Psychoeducation
		Individual and group stress inoculation
Decrease PTSD symptoms, as measured by the PTSD Checklist—Civilian Version (PCL-C <50) within 12 months	Self-report PCL-C	Coordinating care with psychiatrist
		Individual and group stress inoculation
		Cognitive therapy
		Exposure therapy
		Seeking Safety
Maintain sobriety from alcohol for duration of treatment	Self-report	Coordinating care with residential treatment program
		Self-help groups/AA
		Relapse prevention
		Stress inoculation
		Seeking Safety
Increase social network (e.g., increase identified safe/helpful friendships, reconnect with family, access support groups)	Self-report	Multiple group therapies
		Self-help groups/AA
		Exposure therapy to decrease shame
Increase psychosocial functioning, as indicated by securing employment and housing	Self-report	Coordinating care with residential treatment program
		Stress inoculation
		Seeking Safety
		Exposure therapy

Note. AA=Alcoholics Anonymous; PTSD=posttraumatic stress disorder.

aging attendance at Alcoholics Anonymous [AA] meetings); engaged in stress inoculation training, helping the veteran to enhance his coping resources (e.g., grounding, affect regulation); and employed cognitive interventions to help Mr. B.H. identify his automatic thoughts and thinking patterns and challenge his cognitive distortions. I also utilized basic interpersonal therapy interventions to build a strong therapeutic alliance and encourage engagement with the veteran. This became especially important during times of increased stress that occurred near the beginning of therapy. For example, during the fourth session the veteran reported that a good friend of his was murdered in a dangerous area after relapsing to alcohol. I used a more process-oriented, nondirective approach in helping the veteran process his feelings of grief, fear, and anger around this incident.

Psychopharmacology

Mr. B.H. was referred to the PTSD team psychiatrist after the second session. The primary complaint of the patient was frequent and intense nightmares related to the trauma in the military. He reported that the nightmares occurred almost every night and were frightening, awakened him from sleep, and were increasing over the last 2 months. He had also been feeling more depressed, although

he did not meet the criteria for major depression. Because he was reluctant to try medication, it was decided to take a focused approach to treating his nightmares with the off-label use of prazosin (listed in the guideline as a secondary treatment option with some limited evidence-base support). Mr. B.H. was taking hydrochlorothiazide for his hypertension and was warned about possible orthostatic hypotension. He was started on prazosin 1 mg qhs for 3 days, then 2 mg qhs for 5 days, then 4 mg qhs. He returned to his psychiatrist 2 weeks later stating his sleep was better, "a solid 3 hours." He was taking only 3 mg qhs and was instructed to increase to 4 mg qhs and then 6 mg qhs. Upon returning another 2 weeks later, he stated that he had felt nauseous and uncomfortable on 6 mg and reduced the dose back to 4 mg. The veteran had noticed a decrease in nightmares and more restful sleep. He did not return for further psychopharmacology appointments but remained at the same dose of prazosin for the remainder of this treatment period.

Stress Management Group

Mr. B.H. began to participate in the stress management for PTSD group after our ninth session of individual therapy and participated in all twelve 75-minute classes. We used an unpublished manualized treatment that is utilized in multiple VA PTSD programs. The group consisted of Mr. B.H. and seven other veterans, all of whom had PTSD symptoms and many of whom also struggled with substance abuse issues. The group always began with a brief check-in, followed by 45 minutes of psychoeducation about stress and stress management techniques. It ended with a 15- to 20-minute experiential stress reduction technique (e.g., diaphragmatic breathing, progressive muscle relaxation, guided imagery, meditation). Veterans were asked to practice these relaxation techniques between sessions and report on their experience in the check-in process. Throughout the course of therapy the veteran participated actively in group and practiced the stress-reduction techniques between sessions. He was able to identify two particular relaxation techniques, diaphragmatic breathing and guided imagery, that worked best for decreasing his anxiety. However, he did not report feeling connected to other group members and overall did not feel that participating in the group reduced his PTSD symptoms.

Evaluation After Phase 1

After Mr. B.H. had participated in 22 individual sessions and 12 stress management group sessions, we decided to reevaluate his symptoms and discuss additional treatment considerations. The veteran indicated that he was sleeping significantly better, averaging 5–6 hours of unbroken sleep per night. In addition, the veteran reported that he felt stronger in his sobriety, now having maintained sobriety from alcohol for 11 months. He described having significantly more coping resources that he utilized to cope with both urges to drink and PTSD symptoms. He was practicing his breathing exercises daily and finding these very helpful in decreasing his overall level of anxiety. However, he stated that he continued to experience significant PTSD symptoms, especially intrusive thoughts/images, social isolation, numbing, hypervigilance, and irritability. He also described continuing to feel socially isolated from others in the treatment program, friends, and family. His PCL-C score at this point in therapy was 62, indicating a decrease from his baseline score of 70 but still a clinically significant level of PTSD symptoms.

I proposed exposure therapy as a next step, and the veteran stated that he felt ready to engage in this treatment. I again contacted both his psychiatrist and his case manager to let them know about the change in the treatment plan and the potential short-term increase in his anxiety when first starting exposure therapy. I also encouraged the veteran to let others in his life know that he might need additional support for the next few months (e.g., roommate at program, sister, AA sponsor). Additionally, a new Seeking Safety group was starting at the same time, and so the veteran decided to participate in this group co-led by myself and another psychology fellow. The veteran reported feeling excited to have a place to continue to discuss and build coping skills.

Phase 2: Exposure Therapy and Seeking Safety

Individual Exposure Therapy

Mr. B.H.'s exposure therapy consisted of 30 individual sessions over a 5-month period, typically meeting twice weekly. Of the individual sessions, 16 were 90-minute sessions in which the veteran participated in imaginal exposure. The other 14 were shorter sessions (typically 60 minutes) where the veteran processed the feelings that arose for him while retelling

the story of his trauma or listening to his trauma tapes. In each of the 16 imaginal exposure sessions the veteran audio taped his trauma retelling and then listened to it once or twice during the following week. As such, the veteran listened to his trauma tapes 30 times during the course of therapy.

The first session was spent providing psychoeducation about exposure therapy and creating a thorough safety plan for potential increases in anxiety and alcohol cravings. I spent time explaining the structure of the therapy and the rationale for exposure. We identified 10 ways of coping with anxiety and urges that might arise during the course of exposure and wrote them down on a "coping card" that he carried with him at all times. We also discussed that he could call me during VA hours if he needed to check in or listen to my voice mail after hours in order to become grounded. Lastly, we identified the primary trauma that would be the focus of exposure therapy. Because most of his nightmares and intrusions were related to his MST, as opposed to his childhood molestation, we decided to focus on this as the primary event for the trauma retelling. Between the first and second sessions the veteran decided to begin the process of exposure by creating a written account of the trauma and its impact on him. The following is an excerpt from his writing, which highlights the strong link between his PTSD and substance use:

> I was so afraid of being asked what was going on with me that I cut all my friends off and became a loner. I would go off alone and drink. I don't know what people thought and I did not care. Nobody was going to find out the truth. Nobody. You don't tell the army you were raped. Not if you're a man.... I could not go home.... I kept stuffing what happened to me by drinking...the more I drank the more I could be who I wanted to be, 'cause I did not want to be me.... I felt at fault for everything bad that happened then and now.... I deserved it 'cause I did not fight back enough.... I drank to forget and to be anyone but me. I had so much shame and guilt that I just could not let anyone know me, because they might find out my secret. I need to know who I am because I have not known since that night.

During the second session of exposure therapy, the veteran spent 55 minutes telling the story of his rape for the first time while being audio taped. Mr. B.H. connected to his affect readily during session and was extremely tearful throughout the entire recounting of the trauma. During much of the retelling he was clenching his body, and he made absolutely no eye contact. In addition, at the most difficult part of the story the veteran became extremely anxious and vomited. During the processing time after the trauma account the veteran described feeling significant sadness, anxiety, fear, and anger. He felt shame about what had happened and for having shared it with me and continued to have difficulty making eye contact. Given his reactions in this first session, it became clear that part of the goal of exposure would be to help the veteran relax his body while recounting the trauma and to move toward more eye contact after the retelling. In addition, we decided to monitor his vomiting and the number of daily showers to ensure that these decreased over time.

During the third through sixth sessions the veteran recounted the trauma two more times and also spent time discussing the strong emotional reactions he was having both in and out of session. He shared that he was experiencing a significant increase in anxiety as well as urges to use alcohol. It therefore became extremely important for me to adjust the exposure therapy protocol and allow additional time for processing his feelings of anxiety, sadness, anger, and shame, and every week to discuss coping strategies for dealing with his urges to use. At the end of the sixth session the veteran completed the PCL-C and his score was 60, indicating that his PTSD symptoms had not yet improved.

In the seventh session (the fourth in-session recounting of the trauma), the veteran remembered and shared an additional piece of information about the trauma that was particularly difficult for him to say aloud. Although he had been able to recount having been raped with objects, this was the first time that he was able to share the whole rape story, which included having been raped by another man. This felt like the first "whole" recounting and the veteran was extremely sad and angry throughout the retelling. We spent additional time at the end of the session ensuring that his anxiety decreased by 50% prior to ending the session and ensuring that we developed an explicit safety plan for him that day.

In sessions 8 through 30 the veteran continued to recount his trauma and to listen to the tapes between sessions. Each week his trauma recounting became easier, and he described progressively less anxiety. Although decreasing the tensing of his body was very hard for him at first, because he reported that he felt

as if clenching was the only "protection" he still had, by the final retelling his whole body was relaxed when recounting the trauma. In addition, after the sixth recounting of the trauma the veteran no longer experienced vomiting and his showering stabilized at three times per day, as opposed to the previous rate of five times per day. In addition, in the final recounting the veteran was able to tell the whole story while also looking at me, representing a significant decrease in his feelings of shame. In the "processing" sessions in which the veteran explored his affective and cognitive reactions to the trauma, Mr. B.H. discussed further the ways that retelling his trauma story also affected his feelings and thoughts toward his childhood molestation. Thus, the processing often occurred for both traumas, even though the recounting only occurred around the MST.

After every in-session and between-session imaginal exposure the veteran filled out a form asking him about his thoughts and feelings. In the processing sessions we would discuss the information on these sheets at length. The following quotes exemplify the progression of his responses during the course of treatment, representing a significant and meaningful shift in his relationship to the trauma:

Question: What are your feelings and thoughts right now?

I feel sorry and sad for that guy on the tape. What those fucking bastards did to him changed his life forever. I feel anger and rage and sick to my stomach. I have to go throw up.

I feel pretty helpless right now. I don't believe that anyone could do that to me. I want them to pay for this.

My feelings are confused. I still have some anger, but not as much. Mostly sadness. I feel alone.

I am not feeling as bad as I have. I know what's coming and know that I have to deal with it.

I must be getting better 'cause my thoughts are pretty clear. What happened, happened. As much as I want to change it, that doesn't change. So, I'm in therapy getting healthy and dealing with it.

I've had a positive week, so right now I feel sadness and empathy, not just for myself but kind of for the assholes that did that to me. How sick they must be.

Question: What did you do that you should be able to feel good about? Can you allow yourself to feel that?

The only thing I feel good about is that I crawled out of there and survived. No!!!

I do feel relief that I'm not running away right now.

I came here and took risks to trust others with this. It's too heavy for me to carry alone.

I feel good that I can listen to the tapes and relive the evil without walking out and killing someone.

I lived and made it back to the U.S. I feel good about where I am at right now in my life. Maybe by going through all that, that's how I got here!!

In addition to these significant shifts in his affective and cognitive reactions to the trauma, Mr. B.H. also experienced a significant decrease in his PTSD symptoms. His PCL-C score at the end of the exposure therapy was 42, representing a significant decrease in PTSD symptoms and falling below the 50-point recommended cutoff for PTSD in veterans. In addition, during the course of exposure therapy the veteran obtained a job, moved into his own apartment, and began reconnecting with family members. He also developed a friendship with a man at the residential treatment community and at the end of treatment decided to share the story of his rape with this male friend. The veteran maintained his sobriety throughout the course of the exposure therapy.

Seeking Safety Group

During the same time that Mr. B.H. participated in the exposure therapy, he also participated in a Seeking Safety group that consisted of the patient and three other veterans. He participated in all 25 sessions of this 60-minute group over a 6-month period. Seeking Safety is a manualized, group-based treatment for individuals with comorbid PTSD and substance abuse/dependence. Each of the groups began with a check-in process and the review of a topic-relevant quote. Veterans then spent 30–45 minutes reviewing psychoeducation/skill-building handouts that are provided and discussing how the topic is relevant to their own experience with PTSD and substance abuse. The group then ends with a checkout in which veterans identify a committed action that

they will take before the next group. Mr. B.H. participated actively in this group and reported that he found the topics extremely beneficial. The most notable difference between his participation in this group and the stress management group is that here, the veteran connected very well with other group members. They would often meet up before group or stay after group to continue the social process. In addition, midway through group, group members shared their telephone numbers with each other and often had out-of-group contact with one another.

Evaluation After Phase 2

After Mr. B.H. completed exposure therapy and while he finished the Seeking Safety group, we met for a final eight sessions to evaluate progress, discuss the course of treatment, identify lessons learned, and process termination. My fellowship was ending, and thus the veteran was going to be transferred to another treatment provider. We discussed at length his multiple areas of progress over the course of the year, including maintaining his sobriety, decreasing his PTSD symptoms, improving his sleep, enhancing his social network, and increasing his psychosocial functioning (e.g., obtaining employment and independent living). We also spent significant time discussing his feelings about the end of the therapeutic relationship, because he felt very connected to me given the emotional intensity of the work that we did together. It represented growth that the veteran was able to process and discuss his feeling of grief about saying goodbye. Lastly, we discussed next steps in therapy, including engaging in in vivo exposure to address a few continued avoidance behaviors and continuing to challenge trauma-related cognitive distortions.

Summary of Guideline Use

Table 29–1 summarizes the guideline recommendations that applied to this patient and whether the recommendations were followed or adjusted. The major complication in applying the VA/DoD treatment guideline to this case was the lack of guidance in treatment options and adjustments for veterans with a long history of alcohol dependence. A particular concern was whether exposure therapy could be employed without increasing the odds of relapse to alcohol. This was addressed by turning to the literature for an evidence-supported treatment (Seeking Safety) as well as noncontrolled clinical observations. Based on these reports from outside the guideline, a treatment plan was established that included the addition of a set of safety parameters while conducting exposure therapy and the concurrent application of Seeking Safety group therapy. These additional interventions allowed for the safe and successful application of exposure therapy.

Although the VA/DoD guideline strongly recommended initiating pharmacological treatment with an SSRI, Mr. B.H. was resistant to trying medications, despite having agreed to a consult. The veteran agreed on a targeted approach with prazosin for nightmares. Thus, patient preference resulted in a deviation from the guideline. The dosing of prazosin was below that in the guideline (which recommends 6–12 mg) because of side effects experienced at doses above 6 mg. Verbal communication with researchers involved in investigating prazosin suggested that lower dosages can be efficacious. Because the targeted symptom of nightmares improved and the veteran was engaged in exposure therapy, no further pharmacological treatments were attempted.

The focal treatment intervention in this case, exposure therapy, also deviated somewhat from that described in the literature because of the veteran's significant history of substance dependence. The number of sessions was able to be titrated in part because of the primary author's more flexible schedule as a trainee. This flexibility is not available in many treatment settings but may have improved the chances of successful treatment response with this veteran.

Ways to Improve Practice

Although this veteran was treated as part of an integrated treatment team, it was still difficult to integrate fully the psychotherapy and pharmacotherapy interventions into a uniform treatment plan. This was in part because of Mr. B.H.'s reluctance to take medications. In an ideal setting, an evidence-based pharmacological intervention would be given that either would work synergistically with the psychotherapy or even pharmacologically enhance the exposure therapy, as has been proposed by some researchers.

As stated, because this case occurred while the primary author was a trainee, there was the ability to provide more sessions than is described generally in the literature. It is possible the same level of effec-

tiveness could have been accomplished with a decreased number of sessions, but because the number of sessions was not a significant limitation in this case, the greater number of sessions increased the probability the interventions could be applied safely.

The primary outcome measure used was the PCL-C. This is a cumbersome measure to apply in a normal clinic setting in which time is a limiting factor. The lack of validated briefer measures of PTSD is a significant limitation in monitoring treatment responses while administering evidence-based treatments. This case could have benefited from more measures aimed at monitoring improvements in social functioning, sense of well-being, and interpersonal interactions.

References

American Psychiatric Association: Practice Guideline for the Treatment of Patients With Acute Stress Disorder and Posttraumatic Stress Disorder. 2004. Available at: http://www.psychiatryonline.com/pracGuide/pracGuideTopic_11.aspx. Accessed March 4, 2009

Brady KT, Dansky BS, Back SE, et al: Exposure therapy in the treatment of PTSD among cocaine-dependent individuals: preliminary findings. J Subst Abuse Treat 21:47–54, 2001

Brady KT, Back SE, Coffey SF: Substance abuse and posttraumatic stress disorder. Curr Dir Psychol Sci 13:206–209, 2004

Davidson PR, Parker KC: Eye movement desensitization and reprocessing (EMDR): a meta-analysis. J Consult Clin Psychol 69:305–316, 2001

Donovan, B, Padin-Rivera E, Kowaliw S: "Transcend": initial outcomes from a posttraumatic stress disorder/substance abuse treatment program. J Trauma Stress 14:757–772, 2001

Foa EB, Davidson JRT, Frances A: The expert consensus guideline series: treatment of posttraumatic stress disorder. J Clin Psychiatry 60 (suppl 16), 1999

Foa EB, Keane TM, Friedman MJ (eds): Effective Treatments for PTSD: Practice Guidelines From the International Society for Traumatic Stress Studies. New York, Guilford, 2000

Institute of Medicine: Treatment of Posttraumatic Stress Disorder: An Assessment of the Evidence. Washington, DC, National Academies Press, 2007

Najavits LM: Seeking Safety: A Treatment Manual for PTSD and Substance Abuse. New York, Guilford, 2002

Najavits LM, Schmitz M, Gotthardt S, et al: Seeking Safety plus exposure therapy: an outcome study on dual diagnosis men. J Psychoactive Drugs 37:425–435, 2005

Triffleman E, Carroll K, Kellogg S: Substance dependence posttraumatic stress disorder therapy: an integrated cognitive-behavioral approach. J Subst Abuse Treat 17:3–14, 1999

VA/DOD Management of Post-Traumatic Stress Working Group: VA/DoD clinical practice guideline for the management of post-traumatic stress. Washington, DC, Veterans Health Administration/Department of Defense, 2004. Available at: http://www.healthquality.va.gov/ptsd/ptsd_full.pdf. Accessed July 18, 2009

This page intentionally left blank

30

Severe, Recurrent Depression Managed in a Remote Setting Via the Internet

An Example of Remote Care Using the HealthSteps System

Dennis Tannenbaum, M.B.Ch.B., FRANZCP
Caroline Spiranovic, Ph.D.

As practitioners, we are aware that despite implementing evidence-based treatment plans with our patients, many will not adhere to treatment and will cease treatment prematurely as a result of intolerable side effects (Nemeroff 2003). Furthermore, large-scale clinical trials of antidepressant medications indicate that remission may be difficult to achieve even in medication-compliant depressed patients (Rush et al. 2006). Also, relapse rates are relatively high (i.e., 67%) in the long term for patients with severe depression even when they are compliant with medications and achieve remission in the acute stage of treatment (Kennedy et al. 2003). To ensure that the maximum number of patients achieve remission and maintain this state in the long term, there need to be systems in place that supplement the work we can physically do as practitioners. These systems would need to allow for regular monitoring of patient compliance and side-effect burden as well as measurement of treatment response.

The case study we present here outlines the use of the HealthSteps (www.healthsteps.net.au) Internet-based system, developed by Sentiens, to deliver care and monitor remotely a depressed patient from a rural and isolated town in Australia. The patient was seen by the first author (D.T., hereafter referred to in the first person) in his private practice in early 2002. The patient agreed to be managed remotely using an emergent version of the e-health system developed by Sentiens. This example is somewhat of a departure from typical practice, in which the patient would be seen via face-to-face consultation, or occasionally via telepsychiatry, and the HealthSteps system would be employed as an adjunct to regular treatment.

To the best of our knowledge, elements of the HealthSteps system developed by Sentiens are already existent in commercial disease management systems in the United States, but the outcome data for these systems is rarely available for public scrutiny.

Setting

Sentiens is a private specialist mental health clinic that operates a busy outpatient clinic and a day hospital with a 52-patient capacity. Sentiens is based in Perth, Western Australia.

Western Australia is Australia's largest state in area, comprising one-third of Australia's total coastline. Despite this, the population size of Western Australia is approximately 2.1 million and therefore accounts only for 10% of Australia's total population. Over 70% of Western Australian's population resides in Perth, the capital city (Australian Bureau of Statistics 2008). Providing medical services for this vast state is difficult, particularly given that a substantial minority of the population is employed in mining in the remote Goldfields region and hence works on a fly in, fly out (FIFO) basis (Australian Bureau of Statistics 2003).

To provide improved health care for the Western Australian population, Sentiens has pioneered an e-health system that can be used by practitioners to manage and monitor patients wherever they may be: close by in Perth, in rural Australia, or working overseas. The resultant HealthSteps system has been implemented in a number of research trials, both in Australia (e.g., Barnes et al. 2007) and the United States (e.g., Harvey et al. 2007), to further refine its features and demonstrate its efficacy in the management of longer-term health conditions.

Illustration

This study illustrates:

- Monitoring and maintenance of progress
- Patient self-empowerment
- Maintaining adherence

Chief Complaint

Mr. P.J. is a 50-year-old man who requests that I manage a recurrence of his depression that has been progressively worsening. The patient completed Sentiens's online assessment battery, which confirmed the telephone history of severe major depression with diurnal mood variation, early morning wakening, daily depression, agitation, and anxiety with two to three low-grade panic attacks occurring weekly. He was finding it hard to motivate himself and was finding each day interminable. He has fleeting suicidal thoughts but no plan or intention to commit suicide.

His general practitioner (GP) had started him on 20 mg/day of paroxetine 3 weeks prior to our consultation, and this had resulted in a small reduction in panic attack intensity. I wrote to his GP to undertake a number of blood tests and perform a physical examination. The test results from the patient's GP confirmed that Mr. P.J. is in good physical health.

Patient History

Mr. P.J. was well known to me because previously I had treated him intermittently for depression over a 10-year period beginning in the early 1990s. Mr. P.J. initially came to see me for assistance in managing what appeared to be a mild episode of depression and to develop strategies to cope with the difficulties he was experiencing in his marriage. His wife, who was in poor health, was becoming progressively more anxious and depressed and was acting in an aggressive manner toward him and the children.

With regard to the patient's history, Mr. P.J. had performed well academically at university, was socially well adjusted, and was the CEO of a successful company. He had a family history of depression, and his mother had been admitted to a psychiatric hospital on several occasions in the past for severe depression that was treated with electroconvulsive therapy (ECT). Just prior to our first consultation, Mr. P.J. had had a very comprehensive recent medical evaluation from his company, and he was deemed to be in very good physical health.

I initially saw Mr. P.J. for three to four sessions of cognitive-behavioral therapy (CBT) before he decided to cease therapy because of an improvement in his situation. However, he returned to see me within 1 month of ceasing therapy because his depressive symptoms had worsened considerably. He indicated he would prefer to persist with CBT because he was resistant to the idea of taking medications. He felt that "taking medication would prove [he] had the family illness" and he wanted to "sort [himself] out without it."

After a further 2 weeks of CBT, it was clear that his condition was worsening and he agreed to try an antidepressant medication. (He initially was started on fluoxetine because this selective serotonin reuptake inhibitor [SSRI] had been recently made available in 1990 and was the first of the SSRIs on

the Australian market.) Fluoxetine was not successful. Mr. P.J.'s condition deteriorated rapidly, and he required inpatient admission.

In the hospital, he was started on dothiepin, a tricyclic antidepressant (TCA) that was still commonly prescribed in Australia during the early 1990s. The patient made a rapid recovery on dothiepin at 150 mg/day and returned to work after 4 weeks.

After 4 months, he ceased the medication without discussing it with me, claiming he was well.

I thereafter saw Mr. P.J. on an intermittent basis for approximately 10 years. During this time he frequently ceased taking his medication without prior consultation and would subsequently relapse.

Other Significant Findings From Assessment

The online assessment battery completed by Mr. P.J. indicated that he was struggling to cope with his day-to-day activities.

DSM-IV-TR Diagnosis

Axis I	Major depression with comorbid panic disorder
Axis II	None
Axis III	None
Axis IV	Work stress
Axis V	GAF score: 50

Treatment Plan Considerations

My current assessment of this patient indicated that the primary problem was major depression with comorbid anxiety associated with low-intensity and low-frequency panic attacks. Although the patient's anxiety and panic symptoms had been relieved somewhat by the paroxetine prescribed 3 weeks earlier by his GP, he was still experiencing severe levels of depression and this was affecting his day-to-day functioning. He had a poor level of functioning (GAF=50) that was affecting his ability to work, socialize, and even self-care.

The patient had recently moved to a small town some distance from Perth and was managing a small business. His poor level of functioning was placing his developing business at risk, and hence the patient recognized he needed treatment. However, he was limited in his ability to leave his workplace for the time needed to visit me in Perth without sustaining significant financial loss because he had no cover for himself in his newly developing business. He was also caring for his now ex-wife, who was chronically ill and unable to care for herself. He requested I manage him by telephone and advise his local GP on treatment.

The primary aim of treatment was to achieve a remission in the patient's symptoms and maintain this progress so that the patient could function well and make a success of his life. The patient previously had a satisfactory response to a TCA, namely dothiepin, but this antidepressant gave rise to side effects that he complained about and was prepared to tolerate only because he was unable to achieve success on another medication. His treatment adherence had not been persuasive, as he had lowered and ceased his medication only to subsequently relapse on a number of occasions. In order to achieve and maintain a remission of symptoms in this patient, it was essential that he be treated with an antidepressant that was associated with a lower side-effect burden. The American Psychiatric Association (APA) released their revised guideline on the treatment of major depression in 2000. This guideline provided ample information regarding the side-effect profiles of newer antidepressants, and there were many newer-generation antidepressants available in Australia that had a lower side-effect profile than TCAs (see Mant et al. 2004).

However, given the patient's prior history of severe and recurrent depressive episodes, he would be considered at high risk for relapse. The patient was going to require regular monitoring as well as early intervention in not only the acute phase of treatment but also in the long term, particularly with regard to managing side-effect burden. Although the APA (2000) practice guideline did not specify a minimum length of time required for continuation and maintenance therapy with patients who had recurrent and severe episodes of depression, there was evidence available at the time to suggest that this patient would benefit from maintenance therapy using pharmacotherapy for at least a 2-year period (Ellis and Smith 2002).

Treatment Goals, Measures, and Methods

The primary aim of treatment was to improve the patient's mood and function with an antidepressant that has a tolerable side-effect burden and provide therapy to help him return to and maintain an optimal

level of function. I discussed the medication options with Mr. P.J., and he agreed to remain on the paroxetine prescribed by his GP in the interim. Paroxetine was an accepted SSRI in the treatment of both major depression and comorbid anxiety and panic disorder (American Psychiatric Association 2000). Moreover, SSRIs are typically associated with a lower side-effect profile than older-generation antidepressants (American Psychiatric Association 2000), and the patient reported in the interim that he was tolerating the side effects. Medications were to be prescribed and monitored by his local GP, who was happy to monitor the patient weekly. I discussed this with the patient's GP via telephone, and she agreed to take primary responsibility for the patient's care.

I discussed the option of using the HealthSteps online program for depression with Mr. P.J., and he agreed to enroll in this program. I wanted my patient to use this program because it would allow me to monitor remotely his progress, provide him with relevant psychoeducational material, and communicate with him via a secure e-consultation facility. The program allows both the patient and practitioner to access and print out a progress report that provides a graph of the patient's progress on the key measures over time and details the patient's medication chart and self-reported medication compliance. This HealthSteps program would measure the patient's progress on three measures developed by Sentiens: the Depression Severity Scale (DSS), the Quality of Life (QOL) measure, and the Side Effects Scale (SES). As opposed to using validated measures that were freely available in the public domain, we chose to construct our own measures for use in HealthSteps to avoid the possibility of future royalty claims or liabilities.

The DSS is a 20-item measure developed by Sentiens to assess severity of depressive symptoms. Respondents rate the frequency with which they have experienced each listed symptom on a 4-point Likert scale (never, sometimes, often, most of the time). From this measure, three subscale scores (general depression symptoms, self-harm, mood symptoms) and a total score may be derived, with higher scores indicative of greater depressive symptoms.

The psychometric properties of the DSS were demonstrated in an unpublished report by Ree in 2002. The DSS was administered to a sample of 169 psychiatric patients of Sentiens clinic, the majority of whom were male (57.4%), with a mean age of 42.17 years (SD=12.35). The DSS demonstrated a high degree of internal consistency, with a Cronbach's alpha for the total scale of 0.94 (n=141). Factor analysis using generalized least squares with oblique rotation supported a three-factor solution: general depression symptoms, self-harm, and mood symptoms. The convergent validity for the DSS was demonstrated using the 42-item version of the Depression Anxiety Stress Scales (DASS; Lovibond and Lovibond 1995) as a benchmark measure of depression, wherein the DSS total score (r=0.86) and DSS mood symptoms scale (r=0.80) had good convergent validity with the DASS depression scale (P<0.01). The DSS self-harm subscale demonstrated low convergence with the DASS depression scale (r=0.54, P<0.01), which was to be expected given that the DASS depression scale does not assess self-harm/suicidal ideation. The discriminant validity of the DSS was not well supported, as moderate correlations were obtained between the DSS total score and the DSS mood symptoms subscale and the DASS anxiety and stress scales (P<0.01). However, there were also moderate to high correlations between the DASS depression scale and the anxiety and stress scales (P<0.01), thus demonstrating the commonality of underlying causes and characteristics of these constructs.

The QOL was designed by Sentiens as a brief screening scale comprising 8 items that measure the degree to which the patient's symptoms are having a negative impact on quality of life. The QOL is used simply to detect changes in impairment associated with symptoms and was not intended as a tool for determining clinically defined levels of functioning. Although the psychometric properties of the QOL have not been established, this measure was designed to be representative of the constructs commonly measured in such scales. Respondents rate the extent to which their symptoms have affected various aspects of their lives on a 4-point Likert scale (not at all, a little, very much so, severely). This measure yields a total score as well as four subscale scores (relationships, risk behaviors, social, work and money) with high scores indicating that symptoms are having a greater negative effect on quality of life.

Finally, Sentiens developed the Side Effects Scale to measure the level of medication side effects experienced by patients completing pharmacotherapy. This scale was designed specifically for descriptive purposes to gauge the level and type of side effects reported, and thus its psychometric properties have

not been evaluated. Respondents rate the extent to which they have been experiencing 30 items relating to medication side-effects on a 4-point Likert scale (not at all, a little, very much so, severely). From the SES, a total score may be derived as well as six subscale scores (akathisia, attention and mood, involuntary movements, parkinsonism, physical difficulties, and sexual problems). Higher scores are indicative of a greater side-effect burden.

With regard to the treatment schedule, the patient was willing to complete monitoring questionnaires on a once-weekly basis for 6 weeks or until a satisfactory response was achieved. The patient also agreed to communicate with me regularly via e-consultation and telephone as necessary. It was agreed that after a satisfactory improvement, the frequency of monitoring questionnaires and e-consultations could be reduced gradually provided the patient remained compliant and continued to improve. The primary aim was to continue treatment with antidepressants until the patient's depressive symptoms were reduced to a normal range and were consequently having minimal impact on his quality of life. I would then maintain the patient with pharmacotherapy for at least 12 months. A secondary aim was to greatly reduce the side-effect burden of the medications he was taking because this would in turn greatly enhance the patient's likelihood of adherence and remission in the long term.

Course

I monitored the patient's progress closely on the DSS, QOL, and SES for a period of 46 weeks. After this time, I primarily monitored and communicated with the patient via e-consultation and telephone for an additional 12 months. The progress graphs for the DSS, QOL, and SES are presented in Figures 30–1, 30–2, and 30–3, respectively. Plotted against the scale scores in each of these three figures is the patient's medication chart, as well as events of interest that arose during our e-consultations. It should be noted that the patient had been taking paroxetine for 3 weeks prior to commencement of this 46-week period of monitoring.

Initially, paroxetine at 40 mg/day did not result in any noticeable change in symptoms (see Figure 30–1). The patient's quality of life was worsening (Figure 30–2), and the SES (Figure 30–3) indicated he was experiencing substantial side effects. An increase in paroxetine dosage to 60 mg resulted in minimal improvement in depressive symptoms (Figure 30–1) and quality of life (Figure 30–2), and he was still scoring in the severe range for depression. During this time, there was a noticeable improvement on the SES (Figure 30–3) that indicated the medication side effects were dissipating.

However, the patient subsequently and abruptly ceased taking the paroxetine without consultation. He reported he was experiencing significant sleep disturbances, which included episodes of nocturnal sleepwalking. When the patient ceased taking the paroxetine there was a noticeable worsening in his quality of life, as demonstrated in Figure 30–2. By this time point, the patient had been taking paroxetine for a total of 8 weeks, having commenced this medication 3 weeks prior to our consultation, and I had now been monitoring him for 5 weeks. The APA (2000) guideline recommends that it is appropriate to undertake a review of the treatment plan and medications if a moderate improvement in symptoms is not observed after 6–8 weeks of pharmacotherapy.

Thus, approximately 8 weeks into pharmacotherapy, the patient was recommenced on paroxetine at 40 mg/day. I decided to continue the patient on this lower dose of paroxetine rather than switch to another antidepressant, because of the effectiveness of this SSRI in treating symptoms of both depression and anxiety. The difficulty was that paroxetine at 40 mg as a monotherapy would not produce a sufficient response for this patient and the patient was experiencing significant sleep disturbances. I was aware that in this instance, where a partial response was obtained, the APA practice guideline would recommend the addition of an augmenting agent such as lithium, thyroid hormone (namely triiodothyronine: T_3), or buspirone (a 5-HT_{1A} receptor antagonist) in addition to the SSRI. However, these augmentation strategies were unlikely to alleviate the insomnia experienced by the patient.

Hence, mirtazapine, a norepinephrine-serotonin modulator, at 45 mg/day, was added for its additive antidepressant effect and to alleviate the patient's problems with sleep. In support of this choice, the current Texas Medication Algorithm for major depression supports the switch to a sedating antidepressant medication such as mirtazapine to address insomnia as a side effect (Suehs et al. 2008). Furthermore, there was also evidence available at the time to

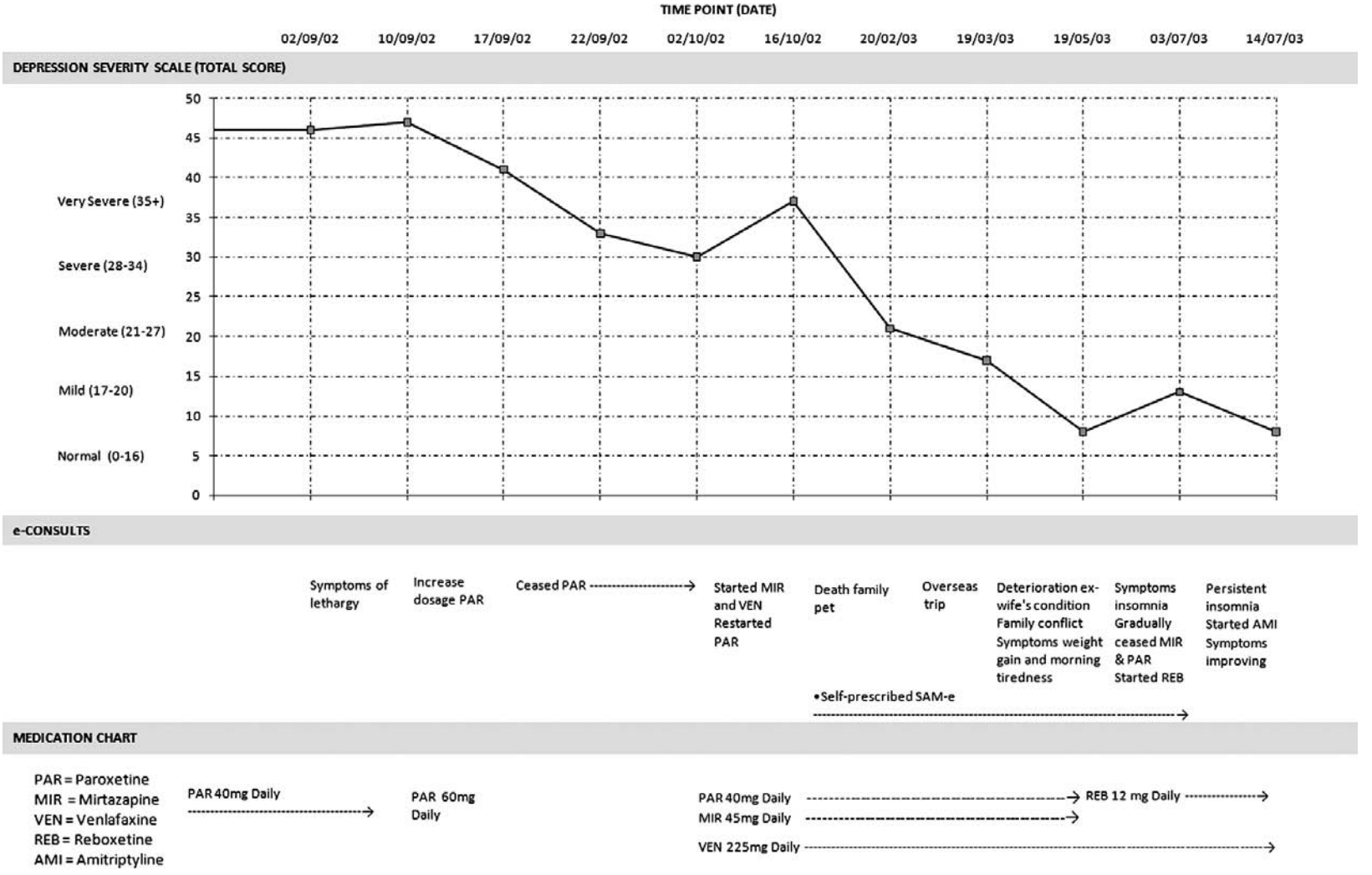


FIGURE 30–1. Total Depression Severity Scale scores, e-consultation events, and medication chart for 46 weeks of treatment.

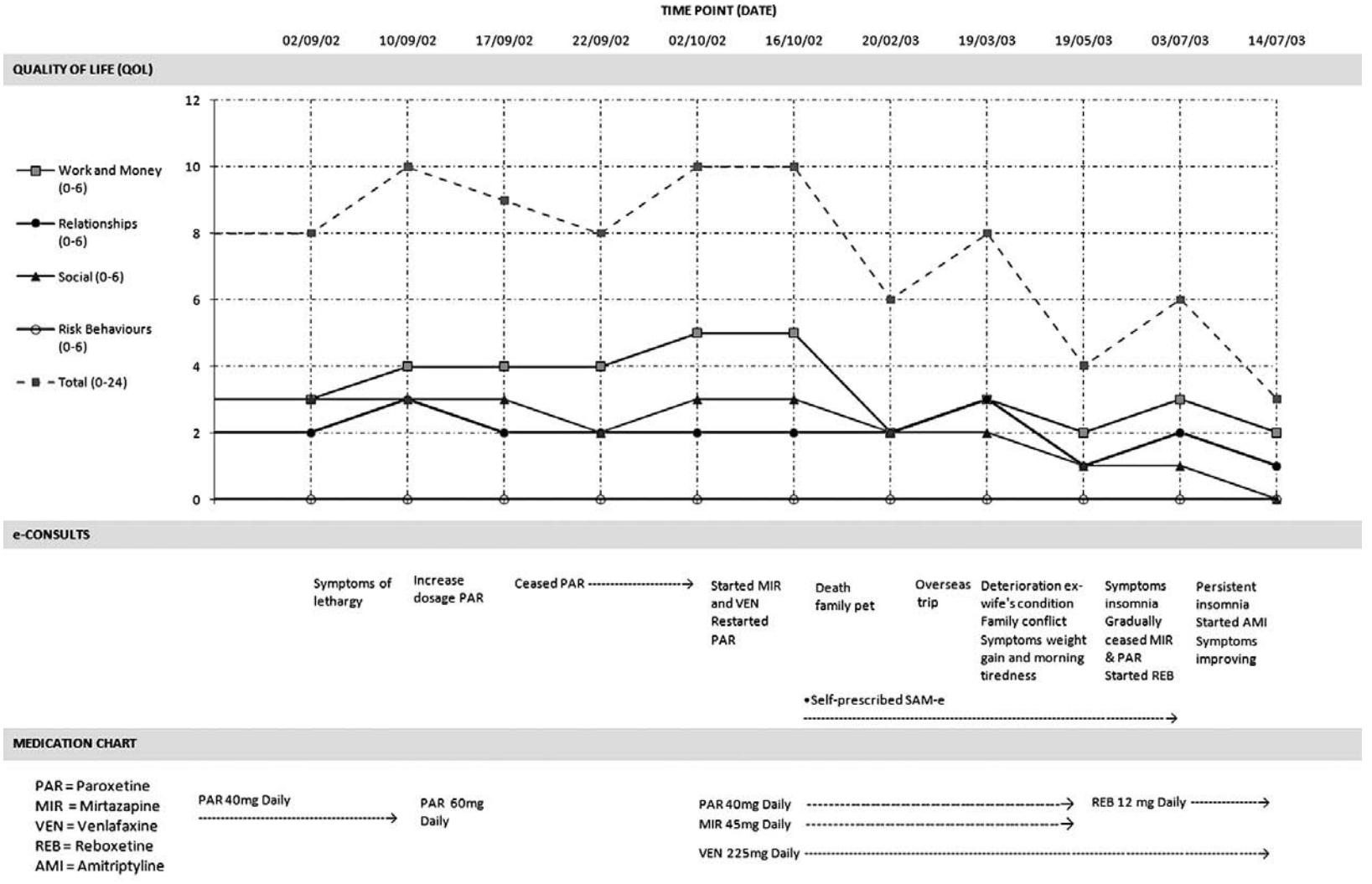


FIGURE 30–2. Quality of Life scores, e-consultation events, and medication chart for 46 weeks of treatment.

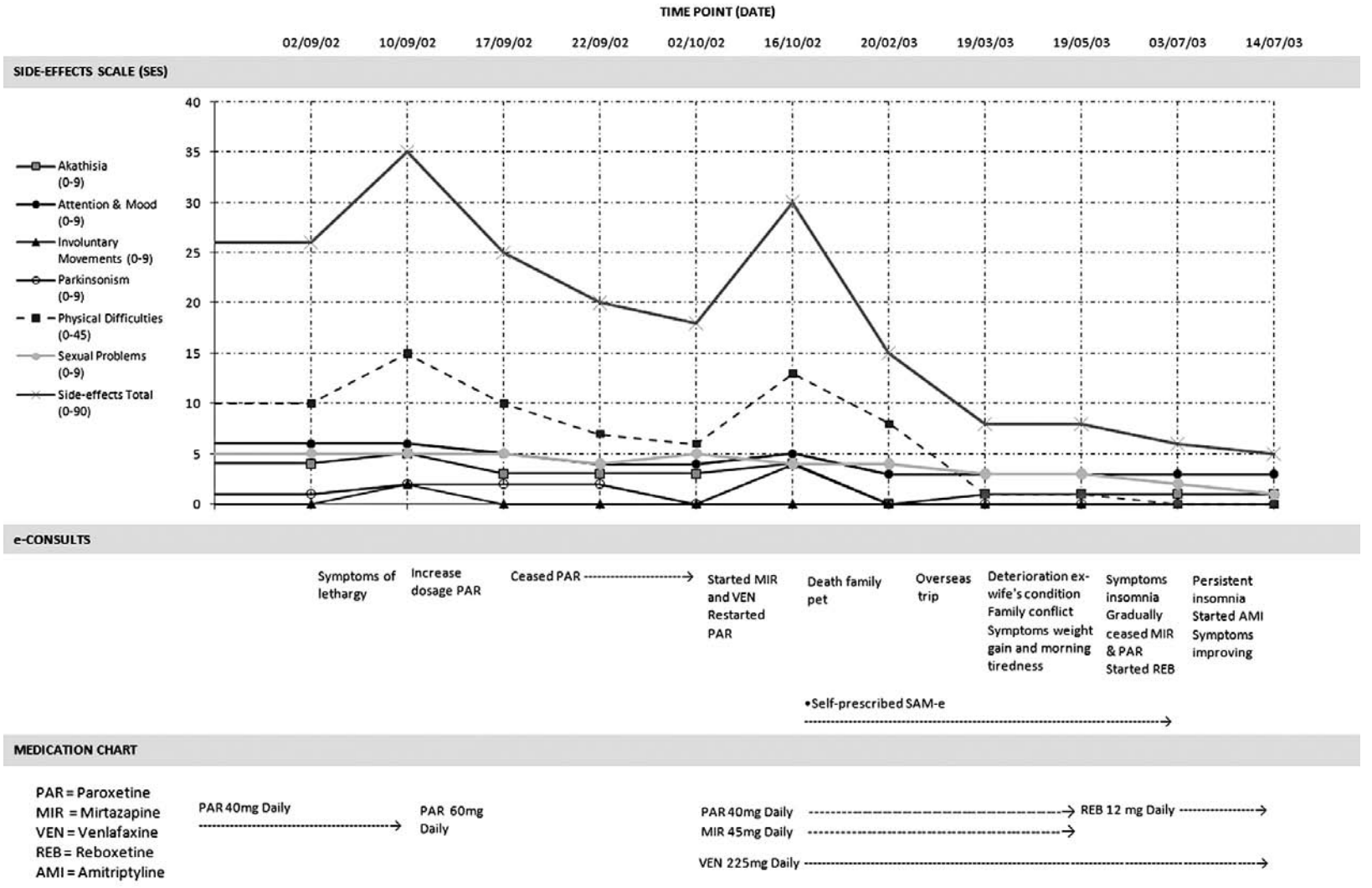


FIGURE 30–3. Side Effects Scale scores, e-consultation events, and medication charts for 46 weeks of treatment.

suggest that the addition of mirtazapine might be particularly helpful given its dual serotonergic and noradrenergic action (Fava et al. 2003).

In addition, venlafaxine, a serotonin-norepinephrine reuptake inhibitor (SNRI), at 225 mg/day, was added because of its strong antidepressant action and suitability as a relatively safe adjunct medication to SSRIs. I was aware that the combination of mirtazapine and venlafaxine was the final step of treatment in the STAR*D trials for those patients who had demonstrated relatively high levels of treatment-resistant depression (Fava et al. 2003). Furthermore, this combination was associated with a lower side-effect profile and higher adherence rates than alternatives for treatment-resistant depression such as tranylcypromine, which is a monoamine oxidase inhibitor (McGrath et al. 2006). I was also aware that mirtazapine and venlafaxine had superior safety indices when compared with MAOIs (Kent 2000). It was therefore reasonable to assume that this powerful combination of antidepressants would be tolerated by the patient and would result in a rapid improvement in symptoms and functioning.

Upon commencing the combination of mirtazapine, venlafaxine, and paroxetine, the patient initially experienced an increase in depressive symptoms (see Figure 30–1), which could largely be attributed to a sudden increase in medication side effects (see Figure 30–3). However, this initial increase in symptoms and side effects was to be expected with this powerful combination of medications. The patient thereafter progressively improved, although progress was slower than anticipated. Within 18 weeks of commencing these new medications, the patient's depressive symptoms (see Figure 30–1) had decreased from the very severe range to within the moderate range. This was a marked improvement, indicating that the trial was in part a success. However, the ultimate aim was to achieve a full remission in symptoms, as indicated by a score on the DSS within the normal range. Coinciding with this reduction in depressive symptoms over the 18 weeks was a marked improvement in the patient's quality of life (see Figure 30–2), particularly with regard to work and money, and the medication side effects had lessened considerably (see Figure 30–3).

Further improvement in symptoms, quality of life, and side effects were gained over the next 3 months, to the point where the patient's depressive symptoms were scored in the normal range. It was interesting to observe that these improvements in the patient's depressive symptoms (see Figure 30–1) occurred at a somewhat tumultuous time during which the patient was experiencing significant psychosocial stressors, such as the death of a family pet, as well as a further deterioration in his ex-wife's condition that was resulting in substantial family conflict. These significant stressors also, surprisingly, had little bearing on the patient's quality of life (see Figure 30–2), as he was able to maintain an acceptable level of functioning during this time. These improvements despite adversity indicated to me that the medications were having the desired effect.

During this time of substantial progress, it is noteworthy that the patient, without consultation, commenced a self-prescribed dosage of *S*-adenosyl-L-methionine (SAMe). At the time of my consultations with this patient, there were no guidelines concerning the use of SAMe as an alternative treatment for major depression. There was relatively sparse and inconclusive evidence regarding the efficacy of this naturally occurring psychotropic substance, and studies were typically limited by methodological issues such as small sample size. However, there was some evidence to suggest that this substance exhibited antidepressant effects and that its use was associated with acceptable levels of safety and tolerability in patients (Agency for Healthcare Research and Quality 2002; Pancheri et al. 2002). Given that the patient believed in the efficacy of this substance and I could see no harm in its use as an adjunctive treatment, I was happy for the patient to remain on SAMe. In fact, although I could not rule out the possibility of a placebo effect, given the inclusive evidence regarding use of SAMe, the patient's recent improvements did appear to coincide with the commencement of SAMe.

It is noteworthy that the patient was willing to continue treatment despite adverse effects until a substantial improvement had been obtained. Once these gains in symptoms and functioning had been achieved, the patient was no longer willing to tolerate the medication side effects. About 30 weeks after commencing the combination of mirtazapine, venlafaxine, and paroxetine, the patient reported he was experiencing substantial difficulties with weight gain and morning tiredness, and hence requested a change in medications.

At that point in time, newer generation antidepressants associated with lower side-effect profiles

than their predecessors had become available in Australia. I instructed the patient to start on reboxetine, a norepinephrine reuptake inhibitor. At the time, reboxetine was a relatively new antidepressant, but I was aware of clinical trials that indicated reboxetine was well tolerated and had similar levels of clinical efficacy to many SSRIs (Kent 2000). However, I was also aware that clinical trials indicated reboxetine was associated with impotence in some male patients (Kent 2000), and my own experiences in treating patients through private practice confirmed this. I discussed the side-effect profile of reboxetine with my patient, who had also read widely on this subject, and he indicated he would prefer to try reboxetine because he was unwilling to tolerate the weight gain associated with mirtazapine.

The reboxetine failed to result in an initial improvement, and the dosage was increased to 12 mg/day. This dose had the desired effect of maintaining remission while also lowering side-effect burden. However, the patient's tolerance lessened somewhat and he experienced worsening symptoms of insomnia. In response, venlafaxine was ceased and amitriptyline, a tricyclic antidepressant, was added for its known sedative properties (Phillips et al. 2000) approximately 46 weeks into treatment. Although the APA (2000) guideline would recommend trazodone in this instance, this antidepressant was not available in Australia. Furthermore, benzodiazepine-related compounds were not an adjunctive option because the patient was previously unable to tolerate such compounds. I largely chose to use amitriptyline over other sedative antidepressants because the patient had previously tolerated tricyclics and did not experience the weight gain often associated with such medications (APA 2000).

Clinically, this case brought to light the importance of tailoring treatments to suit individual patients' needs and, in particular, managing intolerances. The combination of reboxetine and amitriptyline appeared to produce sufficiently strong antidepressant effects and also side effects that were below threshold levels of tolerance. This combination was successful in the long term, as the patient continued on reboxetine at 12 mg/day and amitriptyline at 50 mg/day for 12 months before ceasing pharmacotherapy. Furthermore, the patient has maintained contact with me since this cessation and has not relapsed, to the best of my knowledge, in over 5 years.

Summary of Guideline Use

Utilizing this online monitoring system, it is apparent that the patient's progress was markedly slower than the APA guideline would suggest (American Psychiatric Association 2000). At approximately 9 months into treatment, the patient achieved remission. Remission was maintained over the following months, and changes in medications appeared to have the desired effect of further lowering side-effect burden. Although progress was slower than expected, research would suggest that remission is more difficult to achieve in patients with severe and recurrent depression (Kennedy et al. 2003).

Current guidelines recommend longer periods of continuation and maintenance therapy for patients with severe and/or recurrent depression because such patients are at an increased risk for relapse (American Psychiatric Association 2000; National Institute for Health and Clinical Excellence 2007; Royal Australian and New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression 2004). In accordance with these guidelines, I maintained the patient on reboxetine and amitriptyline for an additional 12 months to preserve remission.

Although I was aware of the APA (2000) guideline for treatment of depression, I also knew that the side-effect burden was going to be an important issue to resolve with this patient. It is all too often the case that patients with high levels of side-effect burden cease taking their medications, despite improvements in symptoms, and consequently relapse (Nemeroff 2003). Hence, an essential goal of the treatment plan was to achieve a minimal level of side-effect burden. The present patient, as with many patients, initially improved upon commencement of psychotherapy with SSRIs. Despite the benefits of these medications, it was difficult to maintain an acceptable balance between efficacy and side effects. In the present case, augmentation of antidepressant medications to below side-effect threshold levels proved to be a worthwhile strategy because remission was achieved and maintained for 12 months while also obtaining an acceptably low side-effect burden.

In order to achieve these long-term gains, I was required at times to diverge from what would be considered an ideal approach. For instance, in my experience as a practitioner, the majority of male patients are unable to tolerate reboxetine. However, this patient was a very intelligent and educated man

who had read widely on the available antidepressants. He believed that the side effects of reboxetine would be better tolerated than the weight gain associated with the current medications he was taking. Although I discussed the issue of efficacy versus side effects of this medication, the patient was willing and motivated to tolerate these effects. As was predicted given available guidelines (American Psychiatric Association 2000), the reboxetine resulted in complaints of insomnia, but these were ameliorated with the amitriptyline. This combination was clearly successful in the long term because the patient was able to tolerate the side effects and maintained a remission in symptoms.

This case clearly illustrates the importance of compromise because for many patients the side effects can outweigh the efficacy of many antidepressant medications, and many patients are only willing to maintain pharmacotherapy if a low level of side-effect burden can be achieved. It would be beneficial for future guidelines to outline the processes a practitioner may follow to minimize side-effect burden and dropout while also improving quality of life and, hence, achieving the best results for a given population.

This case study also highlights the importance of the capacity to monitor patient progress, as outlined in the guidelines developed for treating depression (American Psychiatric Association 2000; National Institute for Health and Clinical Excellence 2007; Royal Australian and New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression 2004). Despite these guidelines, practitioners typically find it practically difficult to monitor patients frequently and in great detail. Through the use of the HealthSteps online system, I was able to respond to my patient's concerns regarding side effects in a timely manner and quickly detect any changes in symptoms or quality of life. I believe this capacity to monitor and intervene early was the key to success with this patient because previous to this my patient often ceased medications without consultation and experienced multiple episodes of relapse. Also, and equally important, the HealthSteps system enabled me to develop rapport and establish a working relationship with my patient in the absence of face-to-face contact. The system enabled us regularly to discuss and review such issues as medication efficacy versus side effects, and together we decided on the best course of action.

Ways to Improve Practice

This case study has demonstrated the utility of online disease management systems in monitoring patients with longer-term health conditions. The HealthSteps system developed by Sentiens was designed as an adjunct tool to monitor and manage large numbers of patients over long periods of time. The system facilitates the evaluation of different strategies to maintain patients in treatment through monitoring of symptoms, quality of life, and side-effect burden. As this case study illustrates, patients typically require frequent monitoring to ensure medication compliance. However, it is often the case that practitioners experience barriers, such as heavy caseloads, that preclude frequent and detailed monitoring of patients. Although the system developed by Sentiens can address many of these barriers, it is evident that practitioners may require additional support to effectively monitor and manage large numbers of patients. To address this issue, recently we have begun to explore the possibilities for utilizing the time and expertise of allied health professionals. Allied health professionals may take on the role of case manager and can facilitate practitioners by regularly monitoring patients using the online system and may contact patients who are not progressing well.

In conclusion, developments in information and communication technologies, as evidenced by the emergent HealthSteps system, offer infrastructures that may be used by practitioners to improve their management of patients. These infrastructures enable us to better monitor and maintain patients so that we may break the iterant cycle of relapse and achieve our long-term goal of remission.

References

Agency for Healthcare Research and Quality: S-Adenosyl-L-methionine for treatment of depression, osteoarthritis, and liver disease (Evidence Report/Technology Assessment, No 64). August 2002. Available at: http://www.ahrq.gov/clinic/epcsums/samesum.htm. Accessed March 5, 2009

American Psychiatric Association: Practice guideline for the treatment of patients with major depressive disorder (revision). Am J Psychiatry 157 (suppl 4):1–45, 2000

Australian Bureau of Statistics: Western Australian Statistical Indicators (cat no 1367.5). June 2003. Available at: http://www.abs.gov.au/ausstats. Accessed March 7, 2009

Australian Bureau of Statistics: Western Australia at a Glance (cat no 1306.5). 2008. Available at: http://www.abs.gov.au/ausstats. Accessed March 7, 2009

Barnes C, Harvey R, Mitchell P, et al: Evaluation of an online relapse prevention program for bipolar disorder: an overview of the aims and methodology of a randomised controlled trial. Disease Management and Health Outcomes 15:215–224, 2007

Ellis PM, Smith DA: Treating depression: the beyondblue guidelines for treating depression in primary care. Med J Aust 176:S77–S83, 2002

Fava M, Rush AJ, Trivedi MH, et al: Background and rationale for the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study. Psychiatr Clin N Am 26:457–494, 2003

Harvey R, Smith M, Abraham N, et al: The Hurricane Choir: remote mental health monitoring of participants in a community based intervention in the post Katrina period. J Health Care Poor Underserved 18:356–361, 2007

Kennedy N, Abbott R, Paykel ES: Remission and recurrence of depression in the maintenance era: long-term outcome in a Cambridge cohort. Psychol Med 33:827–838, 2003

Kent JM: SNaRIs, NaSSAs, and NaRIs: new agents for the treatment of depression. Lancet 355:911–918, 2000

Lovibond SH, Lovibond PF: Manual for the Depression Anxiety Stress Scales, 2nd Edition. Sydney, Psychology Foundation, 1995

Mant A, Rendle VA, Hall WD, et al: Making new choices about antidepressants in Australia: the long view 1975–2002. Med J Aust 181:S21–S24, 2004

McGrath PJ, Stewart JW, Fava M, et al: Tranylcypromine versus venlafaxine plus mirtazapine following three failed antidepressant medication trials for depression: a STAR*D report. Am J Psychiatry 163:1531–1541, 2006

National Institute for Health and Clinical Excellence: Depression (amended): management of depression in primary and secondary care (clinical guideline 23). April 2007. Available at: http://www.nice.org.uk/guidance/CG23. Accessed March 7, 2009

Nemeroff CB: Improving antidepressant adherence. J Clin Psychiatry 64:S25–S30, 2003

Pancheri P, Scapicchio P, Delle Chiaie R: A double-blind, randomized parallel-group, efficacy and safety study of intramuscular S-adenosyl-L-methionine 1,4-butanedisulphonate (SAMe) versus imipramine in patients with major depressive disorder. Int J Neuropsychopharmacol 5:287–294, 2002

Phillips MA, Bitsios P, Szabadi E, et al: Comparison of the antidepressants reboxetine, fluvoxamine and amitriptyline upon spontaneous papillary fluctuations in healthy human volunteers. Psychopharmacology 149:72–76, 2000

Royal Australian and New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression: Australian and New Zealand clinical practice guidelines for the treatment of depression. Am J Psychiatry 38:389–407, 2004

Rush AJ, Trivedi MH, Wisniewski SR, et al: Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 163:1905–1917, 2006

Suehs B, Argo TR, Bendele SD, et al: Texas Medication Algorithm Project Procedural Manual: major depressive disorder algorithms. Austin, Texas Department of State Health Services, 2008. Available at: http://www.dshs.state.tx.us/mhprograms/pdf/TIMA_MDD_MDDAlgoOnly_080608.pdf. Accessed November 3, 2008

31

Postpartum Depression Treated in Private Practice

Matthew H. May, M.D.
Margaret F. Reynolds-May, B.A.

Setting

Dr. Matthew May operates a private practice in Menlo Park, California, that consists mainly of self-paying or insured individuals presenting with depression, anxiety, interpersonal, and/or transitional life issues. He combines psychopharmacology and psychotherapy based on the TEAM (Testing, Empathy, Agenda Setting, Methods) framework created by Dr. David Burns. TEAM is a structured treatment methodology incorporating many novel techniques that can be combined creatively to address a variety of mood and relationship problems, addictions, and unwanted habits. Emphasis is placed on testing and information gathering, building and sustaining rapport, collaboration at the motivational level, as well as classic cognitive and behavioral strategies. All techniques and excerpts from publications by Dr. Burns are used with permission.

Illustration

This study illustrates:

- Multimodal treatment of severe postpartum depression using cognitive-behavioral therapy (CBT), electroconvulsive therapy (ECT), and medication management
- Guidelines for assessment and treatment of postpartum depression
- Symptom assessment and monitoring
- Agenda setting in CBT
- Relapse prevention

Chief Complaint

Jane is a well-educated, married, professionally successful 26-year-old Caucasian woman with recurrent major depressive episodes.

Three weeks following the birth of her first child, the patient experienced dysphoria, reduced energy, and crying spells. She interpreted her symptoms as "baby blues" and intended to discuss the problem with her obstetrician. Unfortunately, she became distracted by her child's crying during the office visit and did not report her symptoms. Aside from weight gain, her physiologic recovery from the delivery was normal, including resumption of her menses and lactation (reducing the concern for Sheehan's syndrome). During the next 4 months postpartum, she developed progressively worsening depression, including intense feelings of hopelessness and worthlessness, anhedonia, difficulty making decisions, prolonged daily crying spells, severe anxiety, reduced energy and libido, and suicidal ideation. This late presentation is not atypical and underscores the importance of screening and early treatment by primary care physicians and mental health professionals (Cho et al. 2008; Cox et al. 1987; Dennis and Hodnett 2007; Peindl et al. 2004; Rychnovsky and Brady 2008).

Present Illness

The patient attempted suicide by suffocation on three separate occasions, all within 1 year of the birth of her child. There were elements of impulsivity and desperation in her attempts. She was hospitalized on each occasion and treated with antidepressants, including high doses of sertraline, escitalopram, bupropion, and duloxetine both alone and in combination. She reported little improvement in her symptoms with these medications.

During her third hospital stay, while being treated with bupropion, escitalopram, and duloxetine, she was observed to have some increased mood lability and impulsivity. The question of bipolarity was raised, and she was tapered off of antidepressants and treated with olanzapine for mood stabilization and clonazepam for anxiety. She responded well to this regimen for several days, although she struggled with somnolence. Modafinil was added to alleviate this problem. She also had difficulty with weight gain and glucose metabolism while on olanzapine and was transitioned to aripiprazole. Her mood continued to decline. Having failed standard antidepressant treatment modalities, the patient was offered ECT. After a total of seven ECT treatments, the patient had significant improvements in mood (Montgomery-Åsberg Depression Rating Scale [MADRS] score fell from 45 to 9 during this 6-week hospitalization). Once stabilized, she was discharged to my outpatient practice with maintenance ECT.

DSM-IV-TR Diagnosis

Axis I	Major depressive disorder, recurrent, severe, with postpartum onset; consider bipolar affective disorder
Axis II	None
Axis III	None
Axis IV	Problems with social environment (child care), problems with primary support group, occupational problem
Axis V	GAF scores: 35–40

Treatment Plan Considerations

Suicide and Infanticide Risk

The patient was clearly at high risk for completing suicide (Appleby et al. 1998) and at some increased risk for infanticide (Taguchi 2007). Reliable early screening for suicide risk, infanticide risk, and psychosis was paramount for the success of this case.

Treatment of Postpartum Depression

There is limited evidence for a variety of modalities in the treatment of postpartum depression, both biological and nonbiological (Dennis 2004; Dennis and Stewart 2004). In light of the limited data, we are forced to rely on expert consensus guidelines (Altshuler et al. 2001). The consensus for initial treatment of nonpsychotic, severe postpartum depression is to combine psychosocial interventions with pharmacotherapy (antidepressants). Preferred antidepressants include sertraline followed by paroxetine (safety during pregnancy and lactation is frequently a concern, as discussed later in this section). In terms of psychotherapy, experts preferred CBT and interpersonal therapy (IPT). In the case of failure to respond to medication or in the presence of psychotic features, ECT was agreed upon as an acceptable treatment modality, although no randomized controlled trials exist (Forray and Ostroff 2007; Rabheru 2001). Most experts agreed that providing support in the home for the mother was appropriate and encouraged including the spouse in the psychotherapy.

In the case presented here, trials of antidepressants had either failed or appeared to pose the threat of a "switch" to mania (Altshuler et al. 1995), hence the choice of ECT and atypical antipsychotics for their mood-stabilizing effects. I decided to use a CBT model developed by Dr. David Burns (TEAM), as it is a flexible model with many interpersonal components. In addition, it incorporates meaningful scales to measure the efficacy of each therapy session and progress over the long-term course of therapy. The inpatient team had arranged for support in the home (a full-time professional nanny), and the patient was to attend a partial hospitalization program.

The specific treatment goals for this patient are outlined in Table 31–1.

Treatment of Depression in Bipolar Disorder

There is good evidence for the use of mood stabilizers, especially lithium, in the treatment of bipolar depression (Baldessarini et al. 2003). Although there is evidence that tricyclic antidepressants may increase the risk of causing mania or hypomania, several studies suggest that treating bipolar depression acutely

TABLE 31–1. Initial treatment goals, measures, and methods

Treatment goal	Measure	Method
Reduction in depressive symptoms and suicidal ideation/suicide attempts	Brief Mood Survey and self-report	TEAM–CBT ECT Psychopharmacology
Reduction in anxiety symptoms	Brief Mood Survey	Exposure (increased time in mothering role) Response prevention (decreasing assistance from child care services) Cognitive restructuring
Increased independence and improved functioning	Self-report, observation	TEAM–CBT
Improvement in relationship	Brief Mood Survey	Patient declined treatment initially

Note. CBT=cognitive-behavioral therapy; ECT=electroconvulsive therapy; TEAM=Testing, Empathy, Agenda Setting, Methods.

(6–12 months) with antidepressants, especially selective serotonin reuptake inhibitors (SSRIs), is less likely to result in a "switch" to mania or hypomania than previously was thought (Gijsman et al. 2004; Leverich et al. 2006; Sachs et al. 2007). Meanwhile, longer-term treatment with antidepressants does seem to promote mood destabilization, especially if there are manic symptoms present (Ghaemi et al. 2008; Goldberg et al. 2007). Whether antidepressants are actually effective in reducing the symptoms of bipolar depression is debatable, especially when manic symptoms are also present (in which case there appears to be an increased risk of mania but no reduction in depressive symptoms) (Gijsman et al. 2004; Goldberg et al. 2007; Sachs et al. 2007).

There are data showing an increased chance of recovering from depressive symptoms and earlier recovery when patients are taking part in intensive psychotherapy (as compared with patients who are not participating in therapy) (Miklowitz and Otto 2007). Therapy focusing on interpersonal problems and sleep patterns has empirical evidence of its effectiveness (Ellen Frank's Interpersonal and Social Rhythm Therapy [Frank 2005]). There are other special considerations in treating bipolar illness, including recognizing prodromal phases, the importance of psychoeducation (especially of the genetic nature of the illness and the chronic course of the illness), and measuring compliance with medication (Leahy 2007).

Treatment of Anxiety

The patient experienced nearly paralyzing anxiety. Anxiety is known to exacerbate depressive symptoms and is a risk factor for suicide (Bolton et al. 2008). I felt it was necessary to measure anxiety regularly throughout the therapy and to address this symptom using a combination of cognitive and behavioral exercises. Many potential sources of anxiety exist for first-time mothers:

- Fears of inadequacy as a mother and worries about making parenting errors
- Financial worries as the household takes on additional expenses and may be losing income (depending on the employment arrangements of the mother)
- Fears of losing friends and social support; feeling pressured to maintain friendships despite exhaustion; feeling self-conscious
- Stressors stemming from problems in the partnership and frustrations in the sexual relationship
- Difficulties acknowledging and expressing (productively) feelings of anger or resentment toward family, friends, and the partner

Because of the etiological complexity of anxiety and other mood problems, one of the most important initial steps was to narrow the focus of the psychotherapy, custom tailoring it to the needs of the patient.

TABLE 31–2. Psychotropic medications in pregnancy and lactation

Medication type	FDA pregnancy risk category	Lactation risk category
Benzodiazepines (e.g., clonazepam, diazepam, lorazepam)	D	L3 (diazepam=L4 if used chronically)
Benzodiazepines for insomnia (e.g., estazolam, flurazepam, temazepam)	X	L2/L3
Nonbenzodiazepine anxiolytics and hypnotics		
Buspirone and zolpidem	B	L2/L3
Eszopiclone and zaleplon	C	
Tricyclic and heterocyclic antidepressants (e.g., amitriptyline, clomipramine, desipramine)	C (maprotiline=B)	L2/L3 (doxepin=L5)
Selective serotonin reuptake inhibitors (e.g., citalopram, fluoxetine, sertraline)	C (paroxetine=D)	L2/L3
Other antidepressants (e.g., bupropion, venlafaxine, mirtazapine, nefazodone)	C (bupropion=B)	L3 (nefazodone=L4)
Atypical antipsychotics (e.g., olanzapine, risperidone, quetiapine)	C (clozapine=B)	L2: olanzapine L3: aripiprazole, clozapine, risperidone L4: quetiapine, ziprasidone

Note. U.S. Food and Drug Administration (FDA) pregnancy risk categories: A=controlled studies show no risk; B=no evidence of risk in humans; C=risk cannot be ruled out; D=positive evidence of risk; X=contraindicated in pregnancy. Lactation risk categories: L1=safest; L2=safer; L3=moderately safe; L4=possibly hazardous; L5=contraindicated.

Source. Adapted from ACOG Practice Bulletin Number 92, April 2008; see original publication for greater detail.

Breastfeeding and Psychotropic Medication

The American College of Obstetricians and Gynecologists (ACOG) regularly updates its safety guidelines for the use of psychotropic medications during pregnancy and lactation. For women with bipolar illness who are pregnant or in the puerperium, the risks and benefits of psychopharmacological treatment must be balanced for the individual and the specific agent (Dodd and Berk 2006). All psychotropic medications enter breast milk, and none have been proved conclusively to be safe to the child. Some of these medications are more safe than others, and there are several options among antidepressants in the "L2" category (with L1 being safest and L5 contraindicated; see Table 31–2) (ACOG Committee on Practice Bulletins—Obstetrics 2008).

Although mothers with depression frequently would prefer not to expose their child to any risk, most physicians concur that depression itself poses a threat to the child (ACOG Committee on Practice Bulletins—Obstetrics 2008; Murray et al. 1996; O'Hara et al. 2000). Mothers with severe depression who are tempted to forego treatment should know that untreated depression is associated with its own risks. In the April 2008 Practice Bulletin, ACOG summarizes this concern concisely:

> Maternal Psychiatric Illness, if inadequately treated or untreated, may result in poor compliance with prenatal care, inadequate nutrition, exposure to additional medication or herbal remedies, increased alcohol and tobacco use, deficits in mother-infant bonding and disruptions within the family environment.

Hence the decision is complicated by the necessity to weigh several options simultaneously: the pros and cons of breastfeeding with or without medication and bottle-feeding with or without medication. This is complicated further by emerging evidence that breastfeeding may be quite beneficial to the child's health (Ip et al. 2007). In this case, the patient decided against breastfeeding, which afforded a wider array of options pharmacologically. Foregoing breastfeeding may be interpreted as a sign of inadequacy in women with depression.

Risk of Future Episodes and Family Planning

Women who have had an episode of postpartum depression appear to be at higher risk of subsequent episodes of postpartum depression (Harlow et al. 2007). Women who are depressed during pregnancy are more likely to have children with lower birth weights, growth retardation, and increased postnatal complications (ACOG Committee on Practice Bulletins 2008; Murray et al. 1996). There is some limited evidence that psychotherapy and psychopharmacology can be employed to reduce this risk, but this must be weighed against the findings that most commonly used antidepressants—SSRIs—have been associated with reduced birth weight for gestational age and increased risk for respiratory distress and cesarean section (Newham et al. 2008; Oberlander et al. 2008; Ramos et al. 2008).

Course

Visit 1

The early visits were 2 hours in duration, twice per week. The patient was enrolled in a partial hospitalization, was undergoing maintenance ECT every other week, and had in-home child support. She had been familiarized with a "gradual exposure" model, in which she would slowly, over the course of several weeks, return to the home and acquire increasing responsibilities as a mother. Presession and postsession scores for depression, anxiety, anger, and relationship satisfaction over the course of therapy, assessed by using using the Brief Mood Survey (Burns 2007), are presented in Figures 31–1 through 31–4.

Testing/Information Gathering

I took a detailed history in the initial visit, facilitated by conversations with her inpatient providers. We paid particular attention to the events surrounding previous depressive episodes and her cognitions at those times as well as her recent depression and suicide attempts.

Initial Conceptualization

When depressed, the patient saw herself as fundamentally inadequate, defective, and flawed. She "compensated" for these perceived deficiencies by being meticulous and cautious and setting very high goals for herself. This behavior had given her certain advantages both academically and professionally, including an impressive attention to detail as well as a tendency to consider and prepare for nearly every potential negative outcome. So long as her performance was acceptable to her, she had a good deal of confidence and self-esteem. However, she felt overwhelmed by the new and unfamiliar responsibilities of motherhood. Whenever she did not live up to her own high standards, she would criticize herself ruthlessly. She would worry excessively about even minor decisions and became convinced that she was a bad mother. Suicidal ideation occurred when she believed this thought so completely that she became convinced that her family would be better off without her. She also felt some pull toward believing she was inadequate and defective because this offered her, both in fantasy and reality, an escape from the responsibilities of motherhood.

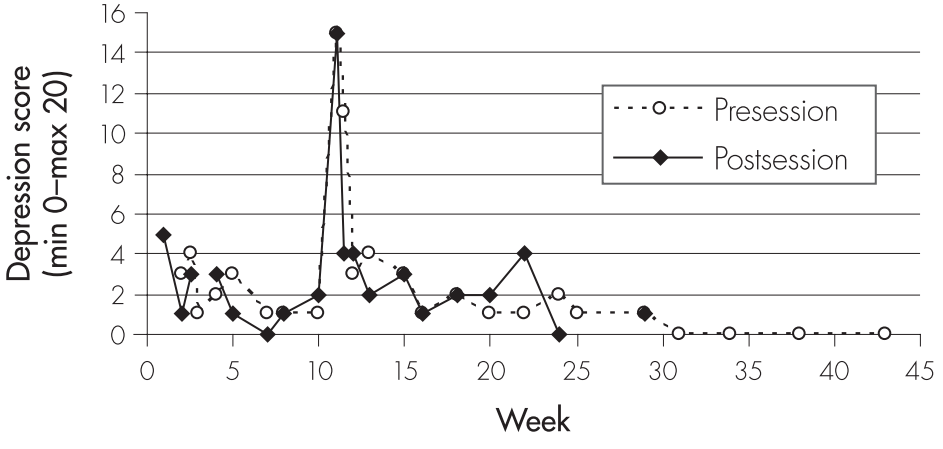


FIGURE 31–1. Brief Mood Survey depression scores over time pre– and post–therapy session.

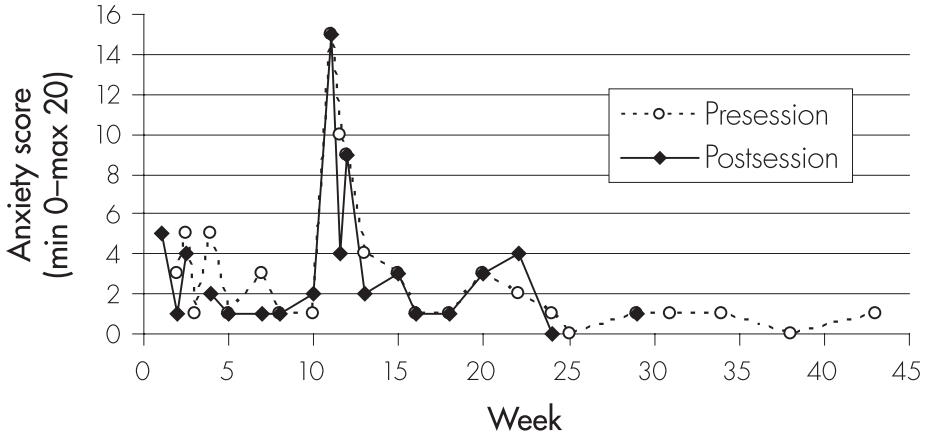


FIGURE 31–2. Brief Mood Survey anxiety scores over time pre– and post–therapy session.

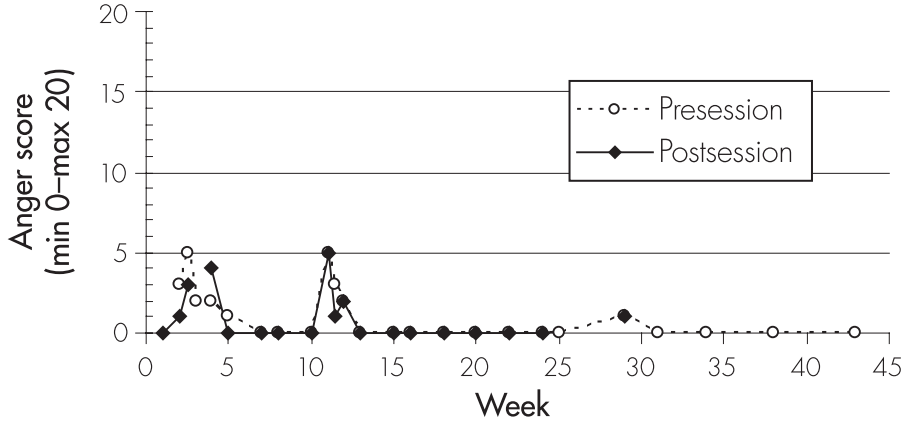


FIGURE 31–3. Brief Mood Survey anger scores over time pre– and post–therapy session.

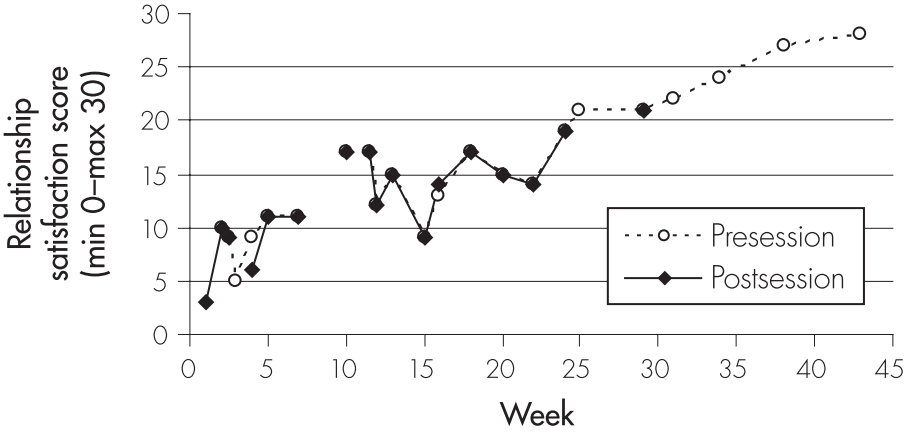


FIGURE 31–4. Brief Mood Survey relationship satisfaction scores over time pre– and post–therapy session.

Empathy

One thing I like about the TEAM model is that if I forget what I am supposed to be doing or get stuck somewhere, I can often simply repeat the acronym and discover where I went wrong. Usually it is somewhere in the middle, either empathy or agenda setting. One of those instances, when I felt like a deer in the headlights, was when my patient said, "I'm a bad mother. My family would be better off without me." This was no cocktail conversation. It will seem easy if you just read along at this point, so in some ways it is better if you imagine what you might say to a patient in that situation or write down what you would say after they say "I'm a bad mother. My family would be better off without me." Of course there would be no right or wrong answer; much is dependent on the relationship with the patient and the setting. I tried to resist the urge to offer advice or reassurance, which is almost never effective and often sounds patronizing. Instead I tried to express my understanding and my own feelings in a curious and respectful manner:

> My heart goes out to you, Jane. You believe you are a bad mother and that your family would be better off without you. I can imagine that you're feeling inadequate and defective as well as discouraged and frustrated. I admire you for telling me that and for how much you care about your family. Do I understand the situation correctly?

Meanwhile, I noted to myself (for later use in agenda setting) that one possible piece of resistance to overcoming depression is that she may believe that her depression is necessary to prevent harm from befalling her family and that recovery would mean greater exposure to the child and hence greater damage to the child. I am not saying this is objectively true, just that it might be true from the patient's perspective.

Safety

I have to admit that this patient made me a little nervous, considering the number of recent suicide attempts she had made. However, she convinced me that she would not attempt suicide again and that she would be accountable for her own safety, including going to the hospital if necessary. She had never had thoughts of harming her child but agreed to share any such thoughts if they were to occur rather than acting on these thoughts.

Visit 2

We discussed her medication regimen, and she expressed an interest in discontinuing aripiprazole because she was noticing increased appetite on this medication and, despite regular exercise, was struggling to lose the weight from her pregnancy. I was concerned about taking her off of a mood-stabilizing agent and suggested adding a different mood stabilizer before discontinuing aripiprazole, to which she agreed. Lithium was perhaps the best choice for mood stabilization, especially in the presence of suicidal ideation. However, with ongoing ECT treatments, this was not an option in the short term (Baldessarini et al. 2006b). I considered lamotrigine to be a reasonable alternative based on evidence that this drug prevented relapse (Calabrese et al. 2003). After discussing risks/benefits and alternatives, she was started on 25 mg/day.

We worked briefly on communication strategies to improve her relationship with her husband and family. This later became a focus of our work, although it will not be described further in this chapter. A gradual improvement in her relationship satisfaction scores can be seen throughout the therapy.

Visit 3

We used a daily mood log to begin to identify her negative cognitions. The situation we identified was a time at home when she, her nanny, and her child were together. She felt 100% inferior and inadequate, 100% hopeless, 100% ashamed, 100% sad, 90% upset and disappointed with herself, and 75% resentful. She had the following thoughts:

- I'm a bad mom.
- I'll always be a bad mom.
- I'm a failure.
- My friends will be disappointed in me.
- I will lose my friends.
- I should never have had a child.
- If I hadn't had a child, I'd still be happy.
- I shouldn't think that.
- I deserve to suffer.

She chose to work on the thought "I should never have had a child" because this resulted in extreme amounts of shame and self-directed anger and resentment.

Decision Point

The patient had identified a disturbing negative thought with which she would like help. CBT is rife with methods to defeat negative cognitions, and I wanted to help the patient. I decided to apply a technique called "identify the distortions." We searched for logical errors in the structure of the negative thought. For example, her thought "I should never have had a child" is using harsh language or "should" statements that are inflicting unnecessary suffering; she is "fortune telling" because only in hindsight could she possibly know that having a child would cause her such suffering; she is applying meaningless labels to herself such as "bad mom"; she is blaming herself for factors outside her control; she is reasoning emotionally by telling herself that because she *feels* bad she *is* bad, and so on. In fact, when we looked closely, we found elements of every cognitive distortion in this thought and its supporting thoughts. Upon identifying the distortions, I switched to an "externalization of voices" model in which I played the role of the patient's negative thoughts, attacking her with these thoughts, and giving her the opportunity to defeat or revise them—with which she did fairly well. I was about to pat myself on the back for being a brilliant therapist when I discovered that she was not feeling any better and was even a little upset with the way the session was going…then I realized I had made a huge agenda-setting error!

To better understand agenda setting, by way of analogy, imagine coming across a friend who is flogging his or her own back with a whip. You might be tempted to point out how illogical this is or try to take the whip away, apply some soothing ointment to the wounds, and so on. This would be the equivalent of applying one of the many different *M*ethods in the TEA*M* model without having the appropriate framework in place. The problem is that intervening without permission can result in resistance. The individual who is self-flagellating must have, in his or her mind, at least one "good reason" to be inflicting such punishment. To avoid forcing the patient to dig in his or her heels when we try to help, we use *A*genda setting (TE*A*M). That means asking whether the patient wants help or just need us to understand what they are feeling. If he or she wants help, rather than chasing our tail talking about every moment the patient had that problem, we narrow the scope to one specific time they were having problems, who was there, what happened, and so on. Agenda setting is also about understanding the type of problem being faced so we can avoid resistance by anticipating and paradoxically aligning ourselves with it. In essence, we become the voice of their resistance. The patient can then choose to argue against his or her own resistance or he or she can agree with us and accept him- or herself proudly.

This is a win-win situation for the therapist. It can sometimes be difficult to imagine the advantages of believing a negative thought, however. Often I simply ask my patient whether he or she can think of any reason to maintain his or her negative thought and make a list with them of advantages to *believing* in the negative thought. Can you imagine any advantages to believing "I should never have had a child"? Here are a few that the patient came up with:

Advantages to Believing "I Should Never Have Had a Child"

- It allows me to recall and fantasize about the "good old days," prior to having a child.
- This will remind me to be more careful in the future so I can avoid such errors. It motivates and facilitates self-improvement.
- It feels morally correct to use such punishing language on myself, since I made an error.
- I can insist on being better, rather than having to accept being an "average mom" or one who makes mistakes.
- It prevents me from being angry with my child or my husband, focuses the blame on me, sparing them.
- If I blame myself and punish myself, other people can't. They will have to take pity on me and help me.
- I can avoid the stress of being a mother and continue to feel the safety of being taken care of by my family.
- If I accept defeat early, then I don't have to try and I can't fail.

Examples of Agenda Setting: How to Discuss Resistance With a Patient

"Jane, I think you're right that this thought is causing some of your shame and hopelessness, and I think we could defeat this thought and lift your spirits today. I'm a little hesitant to do that, however, because I've noticed that this thought reflects

positively on you as a person. You've chosen to blame yourself rather than be upset with your husband or your child for all that has happened to you. This thought is an example of your love and willingness to self-sacrifice. Why change this?"

"If we were to defeat this thought and your depression went away entirely, would you be concerned about what others might expect of you? Wouldn't you be even more stressed than you are now with the responsibilities and demands of motherhood?"

"Are you certain that you would be willing to feel good about yourself considering the fact that you sometimes regret having had a child? Isn't it appropriate for you to feel ashamed of thoughts like that? What would be the moral basis of our eliminating your shame?"

Visit 4

I realized my error in the latter part of visit 3 and dropped back to empathy and agenda setting. After a few agenda-setting questions, the patient was ready to apply techniques and motivated to overcome her self-defeating thoughts, which we did in that session and the following sessions.

Methods

The reason we have over 50 techniques in CBT is that we often need to fail at 5 or 10 or 15 before coming upon the technique that is effective. In this case we used many different techniques in rapid succession, each with the understanding that they were unlikely to be effective and we were eliminating options until we could find the right tool for her. For example, we tried identifying the distortions with a straightforward technique, a downward arrow technique, a cost-benefit analysis technique, and so on.

The most helpful technique ended up being a role-play called the "double standard," which allowed her to be a "good mother" to herself. In this exercise, I pretended to be a clone of her, someone who was exactly the same in every way and had experienced everything she had, made all the same decisions, and felt and thought the same as well. I told my story, how I'd been through a painful and difficult pregnancy, how I'd developed a deep depression that required hospitalizations, and how I was sure I had made a mistake by having a child. The patient readily spoke to me (her clone) in an incredibly supportive way. She would say, "No, you're a fantastic mom, you've just been through a lot and you can't expect to be perfect." I would challenge this, saying, "I'm feeling a bit better, but are you just telling me what I want to hear?" This was our first breakthrough in therapy, and her mood improved significantly over the course of those two sessions. I stored away this technique, reminding myself that what works once will probably work again.

Visit 5

Focus shifted away from her mood, which was increasingly less of an issue for her, and onto her relationship, especially tension between herself and her husband. She seemed to expect that her husband should be doing more in the relationship to be kind to and respectful of her, that he should be less critical of her and be more respectful toward her family. She was disappointed and angry with him but reluctant to express these emotions because of a conflict phobia (fear that expressing her feelings would make the problem worse and there would be no solution). She was unwilling to work individually on this problem because she felt he should be doing more of the changing. She agreed to have her husband come in to the next visit.

Visit 6

As planned, her husband came to this visit. He spoke with me individually, giving his perspective on Jane's illness. He had noticed her perfectionism and negative self-talk when goals were not met. He wanted Jane to express more caring and compassion for him. He felt betrayed and thought that Jane had favored her family over him.

Visit 7

Between sessions the patient experienced a complication with ECT, possibly due to insufficient paralysis. She described unilateral myalgias that were so intense that she could not walk normally. The patient's pain had not responded entirely to nonsteroidal anti-inflammatory drugs. I called in a prescription for four tablets of Vicodin and four tablets of clonazepam. The patient responded well and was feeling better by the time I saw her. We discussed ECT during this visit and she said she disliked going back to the hospital because it was a painful reminder of where she had been and the severity of her illness. She was afraid to discuss her discontent with the physician providing ECT, however.

Visit 8

Jane was essentially euthymic at this point. She was very happy to be home full-time after completing day treatment, although she was still disappointed with her husband. We decided that we could not work on that problem because she felt it was mostly his fault. She had selected a new nanny and had some negative thoughts that she brought in to work on, especially the thought that, because she had to have a nanny, this meant she was defective and inadequate. We did some brief agenda setting and then tried a few techniques. Not surprisingly, it was the double standard technique that was able to beat the thought. Her mood improved immediately in session.

Visit 9

Jane was doing well, but one day after increasing her dosage of lamictal to 100 mg/day, she developed a rash on her arms. There was no mucosal involvement and no lymphadenopathy, myalgia, fever, chills, or lethargy. She was aware of what to look for in terms of Stevens-Johnson rash, and we decided to continue with lamictal at the 100 mg dose and see if her body would adjust to the drug. She was motivated by the fact that she was still having trouble losing weight (and therefore did not want to revert to aripiprazole, which she felt had increased her appetite). We again worked on her negative thoughts; as she engaged more and more in the role of mother, she developed new concerns and new negative thoughts. After we wrote those thoughts down, we again applied the double standard technique:

Example of Double Standard Technique

Negative thought (expressed by the therapist):

"I should have been around and not so fearful. Then my child would be running to me."

"I'm her mom. My child should be just as bonded if not more to me because I carried her inside me for so long and I breastfed her, and above all else she should be bonded to me the most."

Positive thought (expressed by the patient):

"You were sick for so long that your child needs some time to be more bonded with you. It just takes more time; it doesn't happen overnight."

Aside from the patient's rash, I was feeling quite confident that this therapy was going to be a success.

Visit 10

Relapse! The patient was clearly deeply depressed again. She was convinced that she was a "horrible mother" and had an assortment of other negative thoughts. Her affect was constricted, and there was psychomotor retardation. She continued to have a rash on her arms that was itchy. There still was no mucosal involvement, but we decided to have her discontinue the drug for the time being. She cried during the session, which she found cathartic, but we did not have time to do any cognitive exercises, and she left feeling horrible. We scheduled another visit in 2 days, the day after her ECT.

Visit 11

The patient was feeling a bit better after receiving ECT the day before. Her depression score had fallen from 15 to 11 on the Brief Mood Survey. We worked hard in this session and added several techniques, including "be specific," "examine the evidence," and "acceptance paradox," to our favorite "double standard." We worked systematically, using a double column with negative thoughts on one side being replaced with positive thoughts on the other (Table 31–3). After this session, her depression score fell to a 4 on the Brief Mood Survey.

Subsequent Visits

It turned out that the patient's depression, when in full force, required several different techniques working in tandem. She used the four techniques described in visit 11 several times on her own and together in session with me to overcome relapses that came up in the next 30 weeks. She became increasingly self-sufficient. She convinced me that she was able to discontinue her medication and did so without event. I monitored her for some time, but her depression continued to improve. She lost the weight and her relationship improved as well. She created an emergency worksheet that she kept in a safe place that contained the list of techniques and how to use them in case of relapse. We did relapse prevention. She decided there wasn't much need for her to visit me anymore, and I told her I would be available in the future if she wanted to meet with me. We terminated and I have not heard from her since.

TABLE 31–3. Negative thought/positive thought and associated methods

Negative thought		Positive thought *(methods)*
You're a bad mother.	→	Some of the time or all of the time? (*be specific* technique)
All of the time.	→	That would be impossible. Even if I tried to be a bad mother all the time I'd accidentally be a good mother occasionally. (*self-defense* technique)
Some of the time, then, you're a bad mother.	→	That's undoubtedly true. When was a time that I was a bad mother? (*acceptance paradox* and *be specific* techniques)
You were a bad mother yesterday.	→	Probably. I forget: What happened? (*acceptance paradox* and *be specific* techniques)
You didn't love your child enough.	→	There's probably some truth in that. One can probably always improve when it comes to expressing one's love, and I'd be grateful for any pointers you have. Perhaps we could see if we could think of three things I did to express my love for my child and three ways in which I could improve. (*acceptance paradox*, *be specific*, and *examine the evidence* techniques)

Summary of Guideline Use

In the absence of a strong body of evidence-based algorithms on best practices in postpartum depression, we are forced to rely on expert consensus guidelines. The most recent consist of a 2005 report, commissioned by the Agency for Healthcare Research and Quality, which systematically reviewed evidence-based practices related to perinatal depression, and a 2001 study by Altshuler et al. reviewing depression in women (Altshuler et al. 2001; Gaynes et al. 2005). In addition, a 2008 publication by a group of nurse practitioners in Ontario formulated 10 evidence-based recommendations for nurses related to screening, treatment, and prevention of postpartum depression (Gaynes et al. 2005; McQueen et al. 2008). Biological and nonbiological treatment modalities for postpartum depression have been systematically reviewed as recently as 2008 (Dennis and Hodrett 2007; Dennis and Stewart 2004). Recommendations from these sources have been pooled and presented in Table 31–4.

One of the key questions is how to identify postpartum depression. Postpartum depression is generally understood as a modifier ("postpartum onset") of a major depressive episode as outlined in DSM-IV-TR (American Psychiatric Association 2000), indicating the episode commenced within 12 weeks of delivery, although some claim this period can last up to 12 months (Forray and Ostroff 2007; Gaynes et al. 2005). Screening is one important method to identify women experiencing postpartum depression but may not be as necessary in a psychiatric setting in which patients are often already referred for a psychiatric concern (versus a primary care, OB/GYN, or pediatric setting). However, measures to evaluate and track progress of symptoms during treatment are still necessary and useful. One screening/tracking measure that has been used extensively in Europe is the Edinburgh Postnatal Depression Scale, consisting of 10 questions regarding mood and symptoms, and using a 4-point Likert scale (none of which are specific to the postpartum period). A number of trials have shown a score greater than 12 on this scale during the postpartum period provided reasonable specificity for postpartum depression (Gaynes et al. 2005). The Postpartum Depression Screening Scale, available at a cost, and the Beck Depression Inventory have also been used in the literature. There exist concerns about the low sensitivity of these screening methods.

Although it remains unclear whether postpartum depression has a specific unique etiology, evaluation of symptoms continues to mirror that of major depression. In my own practice, I have found Dr. Burns's TEAM approach to be very valuable in identifying, measuring, and treating mood and anxiety disorders. In addition to a comprehensive initial screening and evaluation form, the Brief Mood Survey and Therapy Evaluation form is used at the

TABLE 31–4. General recommendations for postpartum depression

Area	Recommendation
Screening and evaluation	
Screening and diagnostic tools	Depending on setting, specific screening tools (Edinburgh Postnatal Depression Scale, Postpartum Depression Screening Scale) or standard depression instrument (e.g., Beck Depression Inventory, Brief Mood Survey) may be useful.
Exclude medical causes	Exclude medical causes for mood disturbance such as thyroid dysfunction, anemia, and Sheehan's syndrome. Initial evaluation should include a thorough history, physical examination, and routine laboratory tests.
Information gathering	If necessary, seek additional information from other providers of patient (e.g., primary care, obstetrics, pediatrics).
Evaluate risk for bipolarity and psychosis	It is important to assess the risk of bipolarity (current and previous symptoms, past manic or hypomanic episodes, family history), and postpartum psychosis to devise treatment plan and avoid "switch" to mania with use of antidepressants.
Evaluate suicidality	Treatment strategy (modalities and setting) will depend in part on risk of suicidality.
Treatment	
Shared decision making	Includes review of both pharmacological and nonpharmacological therapies and discussion that inadequate treatment increases the risk of morbidity in both mother and infant. Preferably occurs before pregnancy commences.
Patient education	Provide information on nature, course, and treatment, including medication use and side effects as well as nonpharmacological treatment options. Use clear, appropriate language.
Devise treatment plan	Work together with patient to devise a clear and acceptable treatment plan (see Treatment Plan Considerations section in text) based on the severity of the symptoms, including setting (e.g., inpatient versus outpatient).
Nonpharmacological treatment options	
Psychotherapy	Individual therapy: CBT and IPT are therapies of choice. Group therapy may also be useful, depending on case. Couples therapy: involvement of spouse may be helpful.
Psychoeducational or support groups	May be especially attractive to patients who feel alone and/or would like to join a community.

TABLE 31–4. General recommendations for postpartum depression *(continued)*

Area	Recommendation
Nonpharmacological treatment options *(continued)*	
Other at-home support measures	If feasible, at-home interventions such as social work visits or household/child care help (either from paid employees or temporary family help) may be appropriate.
Other self-help measures (e.g., bibliotherapy)	Bibliotherapy based on CBT (e.g., David Burns's book *Feeling Good*) should be offered.
ECT	ECT should be considered in cases of severe postpartum depression, especially those with active suicidal ideation. ECT is considered rapid and safe and can be effective.
Light therapy	Two case studies suggest that incorporating a 100,000-lux box for 60 minutes daily beginning with 10 minutes of awakening for at least 3–5 weeks may be helpful.
Medication options	Use standard antidepressant dosages. Typically symptoms diminish in 2–4 weeks.
Step 1: SSRIs are medications of first choice	Consider medications in terms of previous treatment response, risk to patient, lactation concerns, weight gain, likelihood of overdose/self-harm, tolerability, patient preference, and cost.
	Limited RCT evidence is available to show benefit of SSRIs; prophylactic sertraline and fluoxetine treatment after onset have shown some efficacy.
Step 2: If an SSRI is not suitable, there is no improvement after a 12-week course, or if symptoms remain or side effects become intolerable, further medication substitutions or additions may be appropriate.	SNRIs such as venlafaxine, duloxetine, and bupropion may be effective. If patient is breastfeeding, nefazodone should be avoided.
	Tricyclic antidepressants such as nortriptyline may be useful for women with sleep disturbances; however, evidence does not indicate robust efficacy. If breastfeeding, avoid doxepin.
	Anxiolytic agents such as lorazepam and clonazepam may be useful as adjunctive treatment in patients with anxiety and/or sleep disturbances. Avoid diazepam in breastfeeding women.
	Small observational trials indicate that transdermal and/or oral estrogen (200 µg 17beta-estradiol plus 10 mg cyclical dydrogesterone) may boost mood, especially with concomitant antidepressant use.

Notes for medication options:

6–12 months of treatment is recommended. For women with recurrent major depression, long-term maintenance treatment with an antidepressant is indicated.

If there is a risk of bipolarity, atypical antidepressants may be more appropriate for mood stabilization; weight gain should be monitored. If patient is breastfeeding, quetiapine and ziprasidone should be avoided.

Women should be informed that all psychotropic medications are secreted into breast milk; concentrations can vary widely. Infant serum blood levels are not typically monitored except for lithium use, which is typically not advised because of case reports of neonatal lithium toxicity.

TABLE 31–4. General recommendations for postpartum depression *(continued)*

Area	Recommendation
Monitoring	
Use of monitoring tools	Self-report questionnaires (e.g., Brief Mood Survey) can be used before and after each therapy session.
	Baseline assessment
	Session assessments
	12-week review
	Record response to and satisfaction with therapist/therapy
Response to evaluation	Therapist should respond to evaluations of performance and therapy direction/goals.
Prevention	
Relapse prevention	Therapy work should address potential methods of relapse prevention and identification.
Early screening and identification	Women at risk for postpartum illness should be identified prior to delivery; use of screening tools (listed previously) may aid identification.
Prophylactic treatment	There is limited and divided evidence both for and against prophylactic pharmacological treatment in postpartum depression; sertraline may be efficacious at preventing relapse during subsequent postpartum periods. Antidepressants may be started during pregnancy (see Table 31–2).

Note. CBT=cognitive-behavioral therapy; ECT=electroconvulsive therapy; IPT=interpersonal therapy; RCT=randomized controlled trial; SNRI=serotonin-norepinephrine reuptake inhibitor; SSRI=selective serotonin reuptake inhibitor.

beginning and end of each therapy session to track efficacy of therapy modalities—both in the immediate sense and also over the long-term course of the therapy. These forms have been used in a number of studies and have undergone both reliability and validity testing where applicable (Burns and Eidelson 1998; Burns and Nolen-Hoeksema 1992; Burns et al. 1994).

As described in this patient's treatment considerations, a combination of medication and psychotherapy modalities, as well as an at-home intervention (full-time nanny) was used. These and other recommendations for postpartum depression are presented in Table 31–4. In general, recommendations emphasize combining pharmacological and psychotherapy treatments to achieve the best results. However, because of the lack of randomized placebo-controlled trials of medications in this specific subpopulation, expert consensus guidelines and practices are limited in the scope of their evidence. For example, only three placebo-controlled trials of antidepressants for postpartum depression have been published. The first showed that prophylactic treatment of postpartum depression with nortriptyline in women at risk did *not* prevent recurrence any greater than placebo (Wisner et al. 2001). Using the same study design, the second trial did show a significant benefit of sertraline in preventing recurrence (Wisner et al. 2004). In yet another controlled trial in which fluoxetine with CBT was compared to placebo with CBT in women diagnosed with postpartum depression, the authors found no difference between fluoxetine with one CBT session and placebo with six CBT sessions. This finding suggests that CBT may be as effective as an SSRI intervention (Appleby et al. 1997). Despite the lack of strong evidence for treatment with antidepressant medications, many physicians still consider this approach a standard of care.

Other treatment modalities are based purely on small observational trials or case studies, such as the use of estrogen therapy alone or in conjunction with antidepressants (Ahokas et al. 1999, 2001; Dennis 2004; Gregoire et al. 1996). Similarly, bright light therapy has been endorsed in two case studies but remains to be examined in a randomized controlled trial (Corral et al. 2000; Dennis 2004; Oren et al. 2002). These should be considered experimental treatment modalities, but because they pose minimal risk to appropriate patients they may prove helpful adjunct options.

Ways to Improve Practice

Considering the history of suicide attempts and the possibility of bipolarity, I was even more interested than usual in the topic of safety (Baldessarini et al. 2006a). Despite this concern, I believe I relied too heavily on information provided by the inpatient team during my initial risk assessment. Not surprisingly to veterans of our profession, it is a sad truth that some patients who are depressed often speak less of suicide and experience a brightening of their affect as the date of discharge approaches, knowing that soon they will be free to kill themselves without interruption. Sadly, I covered only a fraction of the information in a structured suicide risk interview, relying instead on "signs." For example, she seemed to have an interest in the future and in scheduling appointments, and she denied suicidal urges at the time of our meeting. To my credit, I did ask, and she agreed to meet with me before acting on any suicidal urges that might develop in the future.

In retrospect, I would have preferred to have used a detailed, structured suicide risk assessment instrument such as that created by Dr. David Burns (Burns 2007). Topics to cover in assessing patients' suicide risk include, but are not limited to, the following:

- History of prior attempts, including whether the attempt(s) were planned or impulsive, whether a note was left and what it said, whether a will was created, how long the individual had been feeling suicidal before the attempt, the situation surrounding the attempt (where they were, who was with them, who said what, what they were drinking), what means were used and some crude sense of how lethal this method might be and how lethal it seemed from their own perspective at the time, any conscious or potential unconscious motive or intent, the reason the attempt did not succeed, and the effects on them physically and psychologically (e.g., did they quit work or file for disability)
- General attitudes toward death and suicide in particular, including a listing of all possible reasons they may want to end their life as well as reasons not to end their life, what they expect the

process to feel like and the moment of death to be like, and their sense of what might come after
- Identification of specific situations in which they would be tempted to end their life
- Current, recent, and lifetime history of depressive symptoms, especially the intensity of hopelessness, anger, and anxiety and patterns of substance use
- Current desire to end their life and what plans they have and whether they intend to act on those plans
- Their attitudes toward being hospitalized, whether they would be willing to go to the hospital if they became suicidal or if this would not be acceptable
- Their attitudes toward communicating suicidality to the provider, including how trustworthy they seem to be, whether they would agree to call if they felt suicidal, whether they would prefer to retain the right to end their life
- Other factors that might place them at higher or lower risk (e.g., primary diagnosis, reports of pain or debilitation, troubles with insomnia, family history of suicide, religious beliefs, significant relationships, future events they look forward to, their support network)
- Would they be willing to make an absolute guarantee of their safety in the future, regardless of how hopeless or desperate they feel?

Along these same lines, my initial intake could have been strengthened by inquiring more about the patient's expectations and motivations. Many patients have unrealistic expectations at the outset of therapy that lurk like hidden icebergs in the North Atlantic. Having experienced my fair share of disappointments in therapy, I now use psychoeducational memos and surveys to assess motivation and the risk of premature termination. The best source that I have found for surveys that provide information about motivation for, resistance to, and attitudes toward therapy that are likely to result in early termination and failure if not addressed early are found in *Therapist's Toolkit*, available from David Burns online at www.feelinggood.com.

An example of a common misconception among patients is the notion that all progress will be confined to the time spent with the therapist, and little change or effort will be required on their part. Meanwhile, there is substantial evidence that individuals who participate in psychotherapy homework are more likely to improve than those who simply talk about their problems (Cowan et al. 2008; Legeron 1991). Unfortunately, there are many good reasons why someone might not want to do psychotherapy homework. Some patients might say "I didn't feel like it," others, "I just knew it wouldn't work," and so on. Before reading the next paragraph, imagine what you would say to a patient who made such a statement.

Most of us would try to convince the patient to do homework. We might try logic and reason, threats of never getting better, begging, and groveling. Most of this will either not work or backfire and result in even greater resistance because the patient will feel that he or she is not accepted as is. Another approach is to agree with them that they have a right *not* to do homework and respect their decision, observing that this might not be a problem they want to work on right now and is there something else they want help with? The Buddhists had it right all along: true power comes from letting go. When we do this, the patient will lead the way.

This has been one of the hardest lessons for me to learn as a therapist, especially a psychiatrist trained in the medical model. I have a strong desire to help those who are suffering and I am sometimes tempted to believe that I can "fix" people without their cooperation or make lasting changes in only an hour per week of talking. It has been humbling to discover otherwise.

One major problem in this patient's therapy had to do with medication management. As I mentioned, I chose lamotrigine because of the evidence available to me that this medication increased the interval between depressive episodes. One flaw with the referenced study on lamotrigine, however, was selection bias. The research, funded by the manufacturer, had included in the treatment group only individuals who had already tolerated and responded to lamotrigine. It was later revealed that they had not published results of five negative studies of lamotrigine for depression (Calabrese et al. 2008).

As I developed more experience with the patient's depression, I came to realize that it might not be primarily chemically mediated. For example, the patient responded much more robustly, when at the height of her depressive episode, to a single 2-hour psychotherapy session than to a session of ECT she received a few days before. In my practice, I have

found that acute improvements in mood are not unusual when I am using the TEAM model of psychotherapy. I have also found that when proper relapse prevention is incorporated future relapses can be avoided or their duration limited to the time necessary for the patient to employ, on their own, the techniques that were helpful initially.

Conclusions

The case presented illustrates the complexity involved in treating one individual who is suffering from severe depression arising in the postpartum period. There is much debate about the etiology of this illness, yet there is little in the way of encouraging data on the efficacy of any one particular treatment or preventative strategy. This is not for a lack of treatment options: medications, electric shock, acupuncture, hormone replacement, light therapy, fish oils, and so on, have all been investigated. An ideal treatment would be safe, rapid, and effective for the patient and in the context of breastfeeding. Psychotherapy would meet many of these criteria, though this option is used less frequently than before among psychiatrists, perhaps because of increased medication options or payment schedules that favor prescribing (Mojtabai and Olfson 2008). Then there are questions of efficacy. Seldom does psychotherapy outperform medication in randomized trials. Meanwhile, the efficacy of medication in the treatment of depression is being called into question (Turner et al. 2008). Clearly we need more effective treatment options for patients with depression, one of the leading causes of morbidity and mortality in the developed world (McKenna et al. 2005). TEAM, a structured treatment framework for psychotherapy, was developed by Dr. David Burns and incorporates a wide array of models and techniques for testing and information gathering; building and sustaining rapport; altering motivation; and facilitating change in mood, relationships, and habits/addictions (Burns 2007). The TEAM approach has frequently surprised the author, especially regarding the rapid rate of improvement in single sessions and the reproducibility of results from one session to the next. Whether it will stand the test of more rigorous testing and comparison remains to be seen.

References

ACOG Committee on Practice Bulletins—Obstetrics: ACOG Practice Bulletin: clinical management guidelines for obstetrician-gynecologists number 92, April 2008 (replaces practice bulletin number 87, November 2007). Use of psychiatric medications during pregnancy and lactation. Obstet Gynecol 111:1001–1020, 2008

Ahokas A, Kaukoranta J, Aito M: Effect of oestradiol on postpartum depression. Psychopharmacology (Berl) 146:108–110, 1999

Ahokas A, Kaukoranta J, Wahlbeck K, et al: Estrogen deficiency in severe postpartum depression: successful treatment with sublingual physiologic 17beta-estradiol: a preliminary study. J Clin Psychiatry 62:332–336, 2001

Altshuler LL, Post RM, Leverich GS, et al: Antidepressant-induced mania and cycle acceleration: a controversy revisited. Am J Psychiatry 152:1130–1138, 1995

Altshuler LL, Cohen LS, Moline ML, et al: The Expert Consensus Guideline Series. Treatment of depression in women. Postgrad Med (March):1–107, 2001

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

Appleby L, Warner R, Whitton A, et al: A controlled study of fluoxetine and cognitive-behavioural counselling in the treatment of postnatal depression. BMJ 314:932–936, 1997

Appleby L, Mortensen PB, Faragher EB: Suicide and other causes of mortality after post-partum psychiatric admission. Br J Psychiatry 173:209–211, 1998

Baldessarini RJ, Tondo L, Hennen J: Lithium treatment and suicide risk in major affective disorders: update and new findings. J Clin Psychiatry 64 (suppl 5):44–52, 2003

Baldessarini RJ, Pompili M, Tondo L: Suicide in bipolar disorder: risks and management. CNS Spectr 11:465–471, 2006a

Baldessarini RJ, Tondo L, Davis P, et al: Decreased risk of suicides and attempts during long-term lithium treatment: a meta-analytic review. Bipolar Disord 8:625–639, 2006b

Bolton JM, Cox BJ, Afifi TO, et al: Anxiety disorders and risk for suicide attempts: findings from the Baltimore Epidemiologic Catchment area follow-up study. Depress Anxiety 25:477–481, 2008

Burns DD: Therapist's Toolkit: Comprehensive Assessment and Treatment Tools for the Mental Health Professional. Philadelphia, PA, D.D. Burns, 2007

Burns DD, Eidelson RJ: Why are depression and anxiety correlated? A test of the tripartite model. J Consult Clin Psychol 66:461–473, 1998

Burns DD, Nolen-Hoeksema S: Therapeutic empathy and recovery from depression in cognitive-behavioral therapy: a structural equation model. J Consult Clin Psychol 60:441–449, 1992

Burns DD, Sayers SL, Moras K: Intimate relationships and depression: is there a causal connection? J Consult Clin Psychol 62:1033–1043, 1994

Calabrese JR, Bowden CL, Sachs G, et al: A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently depressed patients with bipolar I disorder. J Clin Psychiatry 64:1013–1024, 2003

Calabrese JR, Huffman RF, White RL, et al: Lamotrigine in the acute treatment of bipolar depression: results of five double-blind, placebo-controlled clinical trials. Bipolar Disord 10:323–333, 2008

Cho HJ, Kwon JH, Lee JJ: Antenatal cognitive-behavioral therapy for prevention of postpartum depression: a pilot study. Yonsei Med J 49:553–562, 2008

Corral M, Kuan A, Kostaras D: Bright light therapy's effect on postpartum depression. Am J Psychiatry 157:303–304, 2000

Cowan MJ, Freedland KE, Burg MM, et al: Predictors of treatment response for depression and inadequate social support—the ENRICHD randomized clinical trial. Psychother Psychosom 77:27–37, 2008

Cox JL, Holden JM, Sagovsky R: Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 150:782–786, 1987

Dennis CL: Treatment of postpartum depression, part 2: a critical review of nonbiological interventions. J Clin Psychiatry 65:1252–1265, 2004

Dennis CL, Hodnett E: Psychosocial and psychological interventions for treating postpartum depression. Cochrane Database Syst Rev (4):CD006116, 2007

Dennis CL, Stewart DE: Treatment of postpartum depression, part 1: a critical review of biological interventions. J Clin Psychiatry 65:1242–1251, 2004

Dodd S, Berk M: The safety of medications for the treatment of bipolar disorder during pregnancy and the puerperium. Curr Drug Saf 1:25–33, 2006

Forray A, Ostroff RB: The use of electroconvulsive therapy in postpartum affective disorders. J ECT 23:188–193, 2007

Frank E: Treating Bipolar Disorder: A Clinician's Guide to Interpersonal and Social Rhythm Therapy. New York, Guilford, 2005

Gaynes BN, Gavin N, Meltzer-Brody S, et al: Perinatal depression: prevalence, screening accuracy and screening outcomes, AHRQ Evidence Report No 119 (AHRQ Publ No 05-E006-2). Rockville, MD, Agency for Healthcare Research and Quality, 2005

Ghaemi SN, Wingo AP, Filkowski MA, et al: Long-term antidepressant treatment in bipolar disorder: meta-analyses of benefits and risks. Acta Psychiatr Scand 118:347–356, 2008

Gijsman HJ, Geddes JR, Rendell JM, et al: Antidepressants for bipolar depression: a systematic review of randomized, controlled trials. Am J Psychiatry 161:1537–1547, 2004

Goldberg JF, Perlis RH, Ghaemi SN, et al: Adjunctive antidepressant use and symptomatic recovery among bipolar depressed patients with concomitant manic symptoms: findings from the STEP-BD. Am J Psychiatry 164:1348–1355, 2007

Gregoire AJ, Kumar R, Everitt B, et al: Transdermal oestrogen for treatment of severe postnatal depression. Lancet 347:930–933, 1996

Harlow BL, Vitonis AF, Sparen P, et al: Incidence of hospitalization for postpartum psychotic and bipolar episodes in women with and without prior pregnancy or prenatal psychiatric hospitalizations. Arch Gen Psychiatry 64:42–48, 2007

Ip S, Chung M, Raman G, et al: Breastfeeding and maternal and infant health outcomes in developed countries. Evid Rep Technol Assess (Full Rep):1–186, 2007

Leahy RL: Bipolar disorder: causes, contexts, and treatments. J Clin Psychol 63:417–424, 2007

Legeron P: [Predictive factors of a good response to cognitive therapies in depression]. Encephale 17:405–409, 1991

Leverich GS, Altshuler LL, Frye MA, et al: Risk of switch in mood polarity to hypomania or mania in patients with bipolar depression during acute and continuation trials of venlafaxine, sertraline, and bupropion as adjuncts to mood stabilizers. Am J Psychiatry 163:232–239, 2006

McKenna MT, Michaud CM, Murray CJ, et al: Assessing the burden of disease in the United States using disability-adjusted life years. Am J Prev Med 28:415–423, 2005

McQueen K, Montgomery P, Lappan-Gracon S, et al: Evidence-based recommendations for depressive symptoms in postpartum women. J Obstet Gynecol Neonatal Nurs 37:127–136, 2008

Miklowitz DJ, Otto MW: Psychosocial interventions for bipolar disorder: a review of literature and introduction of the systematic treatment enhancement program. Psychopharmacol Bull 40:116–131, 2007

Mojtabai R, Olfson M: National trends in psychotherapy by office-based psychiatrists. Arch Gen Psychiatry 65:962–970, 2008

Murray L, Fiori-Cowley A, Hooper R, et al: The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. Child Dev 67:2512–2526, 1996

Newham JJ, Thomas SH, MacRitchie K, et al: Birth weight of infants after maternal exposure to typical and atypical antipsychotics: prospective comparison study. Br J Psychiatry 192:333–337, 2008

Oberlander TF, Warburton W, Misri S, et al: Effects of timing and duration of gestational exposure to serotonin reuptake inhibitor antidepressants: population-based study. Br J Psychiatry 192:338–343, 2008

O'Hara MW, Stuart S, Gorman LL, et al: Efficacy of interpersonal psychotherapy for postpartum depression. Arch Gen Psychiatry 57:1039–1045, 2000

Oren DA, Wisner KL, Spinelli M, et al: An open trial of morning light therapy for treatment of antepartum depression. Am J Psychiatry 159:666–669, 2002

Peindl KS, Wisner KL, Hanusa BH: Identifying depression in the first postpartum year: guidelines for office-based screening and referral. J Affect Disord 80:37–44, 2004

Rabheru K: The use of electroconvulsive therapy in special patient populations. Can J Psychiatry 46:710–719, 2001

Ramos E, St-Andre M, Rey E, et al: Duration of antidepressant use during pregnancy and risk of major congenital malformations. Br J Psychiatry 192:344–350, 2008

Rychnovsky JD, Brady MA: Choosing a postpartum depression screening instrument for your pediatric practice. J Pediatr Health Care 22:64–67, 2008

Sachs GS, Nierenberg AA, Calabrese JR, et al: Effectiveness of adjunctive antidepressant treatment for bipolar depression. N Engl J Med 356:1711–1722, 2007

Taguchi H: [Maternal filicide in Japan: analyses of 96 cases and future directions for prevention] (in Japanese). Seishin Shinkeigaku Zasshi 109:110–127, 2007

Turner EH, Matthews AM, Linardatos E, et al: Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med 358:252–260, 2008

Wisner KL, Perel JM, Peindl KS, et al: Prevention of recurrent postpartum depression: a randomized clinical trial. J Clin Psychiatry 62:82–86, 2001

Wisner KL, Perel JM, Peindl KS, et al: Prevention of postpartum depression: a pilot randomized clinical trial. Am J Psychiatry 161:1290–1292, 2004

This page intentionally left blank

32

Bipolar Disorder Treated in the Kaiser Permanente Health Care System

Richard E. Shanteau, M.D.

Setting

The patient was seen in an outpatient clinic located in the South San Francisco Kaiser Permanente Medical Center. Kaiser Permanente is an integrated managed care organization. Hospital stays were in local contracted psychiatric units. The outpatient clinic offers crisis intervention, short-term individual therapy, and ongoing medication management. Numerous types of group therapies are offered, including day treatment group programs.

Illustration

- Application of guidelines to bipolar disorder
- Monitoring of symptoms
- Use of multidisciplinary team in HMO setting
- Dealing with relapse

Chief Complaint

Mr. E.S. is a 45-year-old tour boat owner/operator who presents with the chief complaint of bipolar manic episode.

Present Illness

This gentleman first came to my office shortly after he had been discharged from a local psychiatric unit where he had spent 2 weeks for treatment of a severe manic episode. He had been on a cruise ship from San Diego to San Francisco when his wife had become alarmed at his agitation and delusional ideas. He was convinced that the ship was full of CIA and Homeland Security agents who were spying on him and whom he was also helping to train. He sent 20 letters to the captain of the ship in 2 days. He had racing thoughts, slept very little for 7 days, was irritable, and was concocting grandiose schemes to get the country straightened out. Three months prior to his admission he had begun to show increasing manic symptoms. He talked of many new inventions that seemed brilliant to him. He gave $10,000 to his daughter for no clear reason and without discussing it with his wife. He bought his wife a new car that she did not want and planned to give his $800,000 warehouse to his boating community. He sold a $30,000 boat for $1. When he revealed his delusional ideas to his wife during the cruise, his family called Kaiser, brought him to the emergency department, and he was placed in the hospital on a Section 5150 legal involuntary psychiatric hold for grave disability.

Other Significant Findings From Assessment

The patient's premorbid personality had been hard-driving and highly productive, with some irritability. An adult son told the outpatient psychiatrist that his

father had been showing manic symptoms for 2 years, with decreased sleep, euphoria, irritability, impulsiveness, anger, and anxiety. He did not use alcohol or drugs, and had never received any psychiatric treatments. His father and paternal grandfather had both committed suicide. His mother had a diagnosis of bipolar disorder and had been on lithium for many years. He had a stable marriage with three children. Medically he only had mild hypertension, treated with hydrochlorothiazide. He was active in a Protestant church and did not use tobacco or caffeine. He was by all measures a successful businessman and citizen.

At the time of admission to the outpatient service he completed the Patient Health Questionnaire–9 (PHQ-9) and scored 8 of a possible 27, placing him in the mildly distressed range. He did not claim any anhedonia or depression, only indicating some sleep, energy, and concentration problems. He denied any suicidality.

DSM-IV-TR Diagnosis

Axis I	Bipolar I disorder, most recent episode manic with psychotic features
Axis II	None
Axis III	Hypertension, mild
Axis IV	Conflicts with wife
Axis V	GAF score (at time of discharge from hospital): 55

Treatment Plan Considerations

Selecting a Guideline

Kaiser Permanente of Northern California promotes the use of the Texas Medication Algorithm Project (TMAP) guideline (Suppes et al. 2005) for medication treatment of bipolar disorder. The more comprehensive American Psychiatric Association (APA) practice guideline (2006) uses nearly identical algorithms for medications for the manic, mixed, and depressed phases of this illness. In the discussion to follow in this chapter, I compare and contrast any differences between the two sets of guidelines as applied to this case.

Manic Episode

Mr. E.S. did not present any diagnostic challenges, coming in with a straightforward history of premorbid hypomanic traits, no drug or alcohol use, a clear-cut family history of bipolar and perhaps other affective disorders, and a full-blown manic episode of psychotic proportion for his first entry into the psychiatric care system. One wonders why the family did not pressure him to come for consultation before the mania reached psychotic levels.

He arrived at my office with the first decisions having already been made by the treating psychiatrist at the hospital. Because he presented with psychotic features, the use of olanzapine and lithium was correct by both guidelines, although the TMAP suggests olanzapine as an alternate stage 1 choice because of safety concerns (stage 1 agents listed by TMAP are lithium, depakote, aripiprazole, quetiapine, risperidone, and ziprasidone; stage 1b choices are olanzapine and carbamazepine, with metabolic and hematologic concerns). The treating psychiatrist noted that his family history included successful treatment with lithium. Lithium levels were monitored several times during his 2-week stay, reaching a maximum of 1.5 mm/L, which prompted a reduction of dose to 900 mg/day. He was started on olanzapine 20 mg/day, which was raised to 40 mg/day after 1 week. Zolpidem 10 mg was given to improve his sleep pattern. The APA guideline mentions short-term use of a benzodiazepine as sometimes useful, and TMAP suggests use of hypnotics for insomnia. His manic and delusional behaviors gradually quieted.

Mr. E.S. had been treated in the past with hydrochlorothiazide for hypertension; this was discontinued because of the interaction of the diuretic with lithium excretion, and he was placed on clonidine, which effectively managed this issue. TMAP suggests clonidine as adjunctive treatment for agitation or aggression as well. He also had some constipation, which was treated with docusate sodium.

Although he remained hyperverbal and not entirely convinced that his beliefs were delusional, he improved enough to be discharged to home and outpatient follow-up.

Posthospital Management in the Kaiser System

Mr. E.S. was now entering the more difficult part of his recovery, that of adapting to new realities and regaining function, hopefully at a level equal to or better than his premorbid functioning. The APA guideline stresses the need to establish a therapeutic alliance, educate the patient and family about the illness, monitor symptoms, ensure safety, promote

awareness of stressors and regular patterns of activity and sleep, and manage functional impairments. In the Kaiser system of psychiatric care, nearly all patients who are discharged from inpatient stays are enrolled in a short course of day treatment, labeled intensive outpatient (IOP). At the South San Francisco Clinic where this gentleman was enrolled, the day treatment meets for 3 hours every Monday, Wednesday, and Friday mornings, in a group room. The format is interactive group didactic lectures and videos, with training given in relaxation techniques, stress and anger management, cognitive-behavioral therapy (CBT), dialectical behavior therapy (DBT), substance abuse issues and recovery, and psychotropic medications. Social workers, psychologists, a registered nurse, and a physician meet with the 8 to 20 patients in various combinations. Each patient is assigned to a staff member who meets with him or her individually at least once a week. The normal course is 2–4 weeks of attendance, sometimes more. During their IOP course the patients meet with the psychiatrist on their case one or more times per week of attendance. This format of care allows for meeting all the recommendations published in the APA's sensible guideline.

Treatment Goals, Measures, and Methods

1. Achieve euthymia within 2 months, as reported by patient and family, scored on the PHQ-9, and observed by treatment staff, using pharmacotherapy, psychoeducation, CBT, and group and family support. The PHQ-9 (Spitzer et al. 1999) is a nine-item instrument used in the Kaiser system for a quick assessment of depression.
2. Return to full functioning in his work within 2 months, measured by self-report of satisfactory performance, using pharmacotherapy and supportive therapy for patient and his wife.

Course

Mr. E.S. presented to IOP and his initial outpatient evaluations as disheveled, fidgety, with loud voice and pressured rambling speech, though he was generally logical. He seemed to understand that he had been delusional but would occasionally offer rationalizations, which hinted that he was unsure. He was cooperative with the treatments outlined, but somewhat reluctant. He attended dutifully, however. He was taking olanzapine 40 mg/day, lithium 900 mg/day, zolpidem 10 mg at bed, and clonidine for his hypertension. Weight was 182 lb (body mass index=28), slightly below his normal weight of 190 lb. My assessment of him was that he was probably always hypomanic, but never truly manic until this episode. I believed he was responding fairly well to the lithium and olanzapine, but likely would need to switch to another neuroleptic soon, because of metabolic issues of weight gain and altered glucose metabolism. His fasting blood sugar was mildly elevated on repeated checks.

He attended IOP every Monday, Wednesday, and Friday mornings as requested, participating actively. I saw him for 10 to 20 minutes individually during most of these mornings, with his wife often present for the interviews with me. He seemed to be clearing up steadily at first, becoming more normal in appearance and speech. He began to complain of feeling weak, which he blamed on the olanzapine. I elected to lower the olanzapine dosage to 30 mg/ day to address this complaint. Serum lithium level was 0.6 mEq/L.

By 2 weeks after discharge from the hospital he reported feeling ready to return to work but continued to complain of side effects, this time complaining of erectile dysfunction. He and his wife met with me to report that they had unilaterally decided to stop the olanzapine 2 days earlier. At that time he looked improved, with no return of manic symptoms, so I reluctantly agreed with their decision, in order to maintain our developing therapeutic alliance. I scheduled a follow-up appointment in 2 weeks.

Unfortunately he deteriorated rapidly, developing suicidal ideation 8 days after stopping the olanzapine. He had also by this time stopped taking the lithium, a medication that he feared. He rarely took any medications, even aspirin. His father had committed suicide after being recommended to take lithium. His wife brought him in for an urgent appointment, and they resumed lithium 900 mg/day at my prompting.

A week later he returned looking very sad, talking of guilt and failure in his career, complaining of not sleeping, feeling hopeless, acting withdrawn, sitting motionless. He denied ongoing suicidality at this visit. I diagnosed a switch into the depressed phase of the illness, and resumed his olanzapine at 20 mg/day. TMAP calls for raising the lithium level to 0.8 mEq/ L or above, continuing the second antimanic agent,

and adding lamotrigine. I delayed starting the lamotrigine because of the earlier good response to lithium and olanzapine, hoping he would respond to resumption of those medications. I planned to see him two or three times a week, so I could change treatments rapidly as indicated.

Second Hospitalization

Two days later he came to the emergency department of the hospital, where the on-call psychiatrist diagnosed a severe depression, near catatonic level. He presented as disheveled, with poor hygiene, psychomotor retardation, blunted depressed affect, and tangential thoughts; he was preoccupied with poor business decisions made while manic and with his current inability to work. The doctor placed him in a psychiatric unit again, on a legal hold for grave disability, potential suicide risk, and treatment noncompliance.

The hospital psychiatrist found an elevated random glucose level of 156. That, combined with his complaints of fatigue and weakness on olanzapine, prompted the doctor to change his olanzapine to ziprasidone 40 mg bid. The doctor also started lamotrigine 25 mg bid for the depression, and continued his lithium. Within 5 days he was noticeably better and was discharged to home.

Return to Outpatient Setting

Week 8

Mr. E.S.'s wife called to report he was very depressed, with little interest in doing anything. She asked for help with ideas on how to lift him out of his depression. She was reminded to bring him in for his appointments the next day. She was noncommittal about family support group sessions. He came in for IOP and psychiatric visits, looking quite depressed, saying "my thoughts are locked up, like my mind has erased itself."

At this point he had been on the lamotrigine for 2 weeks. He had come out of the hospital on ziprasidone 40 mg bid, but it had been lowered to once daily by the on-call psychiatrist when he called in and complained of feeling too tired. He was taking zolpidem at the 20 mg level and was free of suicidal ideation and delusions. When he came to my office I let the lower ziprasidone level remain and increased his lamotrigine to 75 mg/day for 2 weeks. Again I felt constrained against loading up this man with standard dosages of medications, given the vocal opposition to medications both the patient and his wife expressed.

Week 9

Over the weekend the patient had done poorly, being observed to stand in one place and stare. His wife thought he was responding to internal stimuli. In the Monday morning group he looked worse, apparently slipping back into catatonic behaviors. He expressed extreme guilt feelings, worrying about exhausting his wife and letting his employees down. He was very pessimistic about getting better. In my office he was moving somewhat stiffly and showed a fine tremor of his hands and tongue, but had no cogwheeling or dyskinetic movements. I had him increase his ziprasidone to 40 mg bid again and added clonazepam 0.5 mg to 1 mg bid-tid for anxiety and sleep. TMAP supports use of benzodiazepines targeted to anxiety and hypnotics for sleep. I also added bupropion 100 mg/day for 2 days, then to increase to 200 mg/day. This is listed as a stage 4 maneuver by TMAP. For stages 2 and 3, TMAP suggests using lithium and lamotrigine plus quetiapine or olanzapine/fluoxetine combination, which I was avoiding because of the metabolic issues here. By Friday, 4 days after making these medication changes, he was beginning to look somewhat better, with better color in his face and describing being more able to take care of some chores at home. He had a determined attitude by then, relaying the impression that he would soldier ahead through his ordeal.

Discussion

This sequence of events shows the importance of not lowering medications too soon because of bothersome side effects. Perhaps we could not have persuaded this couple to continue the dosages he was on when he came out of the hospital, but he probably would have had a smoother course of recovery from his psychotic episode if the olanzapine and lithium had not been lowered and stopped. He might have slipped into depression anyway, but one suspects that the second hospital stay could have been avoided. And when he did end up hospitalized again, the ziprasidone was lowered almost immediately after discharge, and he suffered an apparent worsening of his depression after that. He responded within 5 days of returning it to 80 mg (40 mg bid), plus having bupropion and clonazepam added.

Week 10

On Monday he still looked and acted fairly depressed. He complained of memory problems and had amnesia for calling a helper to meet him at 6 A.M. He had set his alarm clock for 6 P.M. His serum lithium level was 0.5 mEq/L on 900 mg, so it was raised to 1,200 mg for 3 days and then to 1,500 mg, seeking a level of 0.8–1.0 mEq/L. He was showing signs of improving in that he was speaking up more in groups. I encouraged him to try to return to a reduced schedule of work. I also had him lower his zolpidem dosage from 20 mg to 10 mg/day.

Two days later he returned to report that he had successfully run his boat tours the day before and was feeling better because of this. His wife reported that he was slowly returning to normal, watching TV, sleeping more normally. He still had a mildly depressed affect but was improving rapidly. I raised his lamotrigine to 50 mg bid and had him switch his zolpidem and clonazepam to prn only. I reduced his clonidine to 0.1 mg bid from tid because his blood pressure readings were normal and we wanted to reduce his overall medication load.

He was able to graduate from IOP that Friday, with obvious progress having been made in 1 week. He even expressed optimism about his future and gratitude toward the staff. He registered for a depression class and agreed to have some individual therapy as well as his psychiatric visits.

Discussion

When Mr. E.S. came out of the hospital the second time, his wife was much more accepting of guidance about the medications. I spent a lot of time in our joint sessions explaining the illness to them and coaching them that we needed to make changes more carefully and slowly over time. I reassured them that it was in my interest as well that he be taking only the necessary number and quantity of medications, to minimize side-effect burden on him and to reduce costs.

Week 14

Mr. E.S. and his wife reported he was doing well, especially when working, which he was doing every other day. No tardive dyskinesia or other side effects were mentioned. He showed no evidence of mania or depression and no further amnestic episodes. He was sleeping 8 hours without zolpidem. He was taking lithium 1,200 mg, ziprasidone 80 mg, bupropion 200 mg, lamotrigine 100 mg, and clonazepam 1.5 mg daily, as well as clonidine 0.1 mg bid.

Week 19

Mr. E.S. continued to do well. His only side effect was minor tremor of the hands. He had gained weight and reached 193 lb, slightly above his usual weight. He did not wish to change any medications at this visit. I recommended he modify his diet to lose weight slowly and see his primary care physician for a checkup.

Week 23

The patient continued to do well. His wife reported he was better tempered now than he was premorbidly, when he was more irritable and angry. He was working 2 days per week for himself, being careful not to overdo things. He would help out a colleague on some other days. He and his wife planned another cruise soon. His internist switched him from clonidine to a lisinopril/hydrochlorothiazide combination. He was taking lithium 1,200 mg, ziprasidone 80 mg, bupropion 300 mg, and lamotrigine 100 mg daily, as well as clonazepam prn.

Week 29

The patient and his wife went on a Caribbean cruise and did well except that he was uncomfortable for a few hours, remembering his manic episode during his previous cruise. He had a mildly elevated fasting blood sugar level, and his lithium level was 0.8 mEq/L. He remained euthymic but complained of erectile dysfunction. In response, I lowered his ziprasidone to 40 mg/day for 4 weeks and then discontinued it. My assumption was that any of the medications could be causing the sexual dysfunction, and the guidelines call for removing the atypical neuroleptics when the patient is stable. Again, I encouraged exercise and weight loss as aids to his overall health and potency.

Week 38

He had been off ziprasidone for 3 weeks and continued to function adequately, but he and his wife both commented that he was not as happy-acting as he had been. He was still having erectile dysfunction and low libido. His sleep was fine, and he was using a treadmill daily. He had lost weight and was now at 180 lb. I had him lower his lithium to 900 mg/day and stop the clonazepam to lower his overall medication burden. I also added trazodone in hopes of

assisting with the sexual issues, using it instead of prn clonazepam for sleep. I e-mailed his internist to ask for clearance to use sildenafil as an aid.

Week 47

PHQ-9=1 (no depression). Again the patient and his wife reported he was doing well, continuing to walk and lose weight, sleeping well, able to socialize comfortably even in large groups. He continued to have erectile dysfunction, with incomplete erections and infrequent intercourse. Serum lithium level was 0.8 mEq/L on 900 mg, weight was down to 172 lb. His affect had improved to be quite relaxed and natural, no longer with a flattened quality. I had him take all his lamotrigine at bedtime and reduced the trazodone to prn use only. He took a sildenafil prescription with him, pending approval by his internist.

Week 55

PHQ-9=2 (no depression). He continued to do well. They had gone on another cruise and reported that the sexual issues had resolved with the use of sildenafil. He discussed some guilt at things that had transpired between himself and his son during his initial hospitalization. His weight and lab values remained stable, within normal limits except for a mildly elevated fasting blood sugar level of 94.

Tables 31–1 and 31–2 list the TMAP algorithms for bipolar disorder, manic and depressed phases, and summarizes how closely they were followed in this case.

Ways to Improve Practice

This case illustrates the danger of lowering therapeutic medications too early and too rapidly in the bipolar individual. In retrospect, I would have spent more time with this patient and his wife in trying to slow down their attempts at withdrawing the medications. Since sexual performance problems proved to be a major concern later, perhaps they were causing some of the early resistance to the medications, and discussing sexual issues earlier would have been helpful.

Also, I will more quickly return to the same levels of medicines that were working before, and not be as hesitant to perhaps cause a few extra side effects. I already move quite slowly and deliberately when withdrawing such people from medications, if they permit me to guide them thusly.

I will also be much more likely to consistently use measurement tools such as the PHQ-9. The Young Mania Rating Scale also would have been useful after his first hospital stay, as he was recovering from the manic episode. The act of completing the scale could serve as an important educational moment for the patient and his wife.

Perhaps most importantly, the work of writing up this case has illustrated to me the value of referencing the published treatment guidelines from time to time in order to reduce uncertainty for the clinician and speed recovery for the patient.

References

American Psychiatric Association: Practice Guidelines for the Treatment of Psychiatric Disorders, Compendium. Washington, DC, American Psychiatric Publishing, 2006

Spitzer RI, Kroenke K, Williams JBW: Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study—Primary Care Evaluation of Mental Disorders, Patient Health Questionnaire. JAMA 282:1737–1744, 1999

Suppes T, Dennehy EB, Hirschfeld M, et al: The Texas Implementation of Medication Algorithms: update to the algorithms for treatment of bipolar disorder. J Clin Psychiatry 66:870–886, 2005

TABLE 32–1. TMAP recommendations for bipolar disorder, currently hypomanic/manic

	Recommendation	Followed
Stage 1	Monotherapy with lithium, valproate, aripiprazole, quetiapine, risperidone, ziprasidone, or Stage 1b, olanzapine or carbamazepine, plus targeted adjunctive treatments	Inpatient physician put him on stage 2 treatment immediately
Stage 2	Two-drug combination of lithium, valproate, or atypical antipsychotic, plus targeted adjunctive treatments such as benzodiazepines, hypnotics	Yes. Lithium plus olanzapine and zolpidem
Stage 3	Try another two-drug combination of lithium, valproate, atypical antipsychotics, carbamazepine, oxcarbazepine, typical antipsychotics, plus targeted adjunctive treatments	Not needed
Stage 4	Electroconvulsive therapy or add clozapine or lithium + valproate or carbamazepine or oxcarbazepine + atypical antipsychotic	Not needed

TABLE 32–2. TMAP algorithm for the treatment of bipolar disorder, currently depressed

	Recommendation	Followed
Stage 1	Antimanic + lamotrigine. If lithium, increase dose to ≥0.8	Patient was on lithium and olanzapine for mania before switching to depressed phase.
Stage 2	Antimanic + lamotrigine + quetiapine *or* olanzapine/fluoxetine combination (OFC)	Lamotrigine was started at the second hospital stay (for depression).
Stage 3	Combination from lithium, lamotrigine, quetiapine, *or* OFC	Ziprasidone replaced olanzapine because of metabolic and other side effects.
Stage 4	Lithium, lamotrigine, quetiapine, OFC, valproate, *or* carbamazepine + SSRI, bupropion, or venlafaxine; *or* ECT	Bupropion was added.
Stage 5	MAOIs, tricyclics, pramipexole, other atypical antipsychotics, oxcarbazepine, other	Clonazepam, zolpidem, and trazodone were used.

Note. ECT = electroconvulsive therapy; MAOIs = monoamine oxidase inhibitors; SSRI = selective serotonin reuptake inhibitor.

This page intentionally left blank

33

Consultation-Liaison Psychiatry

Anthony J. Mascola, M.D.

Setting

The psychiatric consultation-liaison service provides psychiatric care and support to patients admitted to the medical and surgical services of the medical center. It is a unique practice environment that provides many opportunities to apply the tenets of the evidence-based medicine (EBM) model. Disturbances of mood and behavior in the setting of serious medical illness might reasonably be considered psychological reactions to the extreme stressors present in such an environment. They could represent exacerbations of a preexisting psychiatric illness. There exists also the possibility of a much broader differential diagnosis of potential causative factors given the high pretest probability of medical, pharmacological, and surgical conditions associated with psychiatric symptoms in these settings. Teasing these factors apart and assigning a diagnosis that might optimally assist the clinician in easing the suffering of the patient can be challenging. The focus of this case is on using an evidence-based approach to diagnosis.

Recommended Prior Knowledge

If you are unfamiliar with the evidence-based diagnostic process, it is highly recommended that you read the introductory chapters in part I of Steven McGee's excellent text on evidence-based physical diagnosis (McGee 2007) or the introductory article by Sackett (1992) before proceeding.

Illustration

This study illustrates:

- Common biases affecting our diagnostic methods
- Importance of the first step in the EBM cycle—assessment:
  - Importance of, and methods for generating, a reasonable differential diagnosis for the patient's symptoms
  - Importance of carefully considering the pretest probabilities of the conditions in the differential diagnosis
  - Making use of the operating characteristics of tests, including those describing diagnostic accuracy (likelihood ratios) and reliability (kappa)
  - Combining pretest probability estimates with likelihood ratios to estimate the posttest probability of a given diagnosis in the differential
- Asking focused, answerable clinical questions to search the medical literature
- Acquiring information
- Appraising information on diagnostic tests for validity and relevance
- Applying information using a shared decision-making model that incorporates published evidence, clinical judgment, and patient and family preferences
- Assessing both patient outcomes and the process of decision making used

Chief Complaint/Reason for Consultation

Mr. C.K. is a 79-year-old married man of Polish American descent admitted to the hospital for surgical removal of a recurrent squamous cell carcinoma of his vocal cord. His surgery occurred 4 days prior to our being consulted to "evaluate and treat the anxiety and depression affecting his postoperative course."

Present Illness and History

From the information available in the chart and the verbal information shared from the consulting physician, Mr. C.K. had a long history of cigarette use, smoking about one pack per day for 45 years. He quit smoking 17 years prior to admission when he was diagnosed with coronary artery disease and had a two-vessel coronary artery bypass graft performed. He was diagnosed as having an in situ squamous cell carcinoma of the right vocal cord 2 years prior to admission and received radiation treatment. He was thought to be in remission until he began noticing increasing hoarseness of his voice. He was evaluated and found to have a recurrent moderately differentiated squamous cell carcinoma of the right vocal cord staged T3 N0 M0. He was readmitted to the hospital for a modified radical neck dissection, supracricoid laryngectomy with cricohyoidopexy, and tracheostomy.

In addition to his coronary artery disease, Mr. C.K. had intermittent atrial fibrillation. He was not taking warfarin. He also suffered from hypertension, hypercholesterolemia, gastroesophageal reflux disease (GERD), and benign prostatic hypertrophy (BPH) and was status post transurethral prostatectomy. Little was known about Mr. C.K.'s past psychiatric, substance use, developmental history, or social history. He was noted at times in the presurgical consultation notes to appear "anxious" when discussing his treatment options; he was known to be married but his current living situation was not specified. There was no mention of his previous education level, employment history, or any past psychiatric diagnoses or treatment. As an outpatient, he was prescribed tamsulosin hydrochloride for BPH, omeprazole for GERD, amlodipine/benazepril for hypertension, atorvastatin for hyperlipidemia, and aspirin for coronary artery disease.

Mr. C.K. was felt to be an appropriate candidate for surgery for his laryngeal tumor despite the multiple other comorbid medical problems that made him a high-risk surgical candidate, because without intervention the tumor would likely continue to expand and cause increasing airway constriction, dyspnea, and eventually suffocation. The procedure was considered an important palliative intervention. He was admitted to the hospital, and the surgical team performed a modified radical neck dissection, supracricoid laryngectomy with cricohyoidopexy, and tracheostomy as planned.

The team noted that the intraoperative course went fairly well. The tumor was removed, and the patient was hemodynamically stable. There were a few developments in the postoperative course, however, that had begun to complicate his recovery, resulting in the team consulting the psychiatry department.

The patient was initially intubated and sedated with midazolam and fentanyl. He received additional intermittent hydromorphone hydrochloride for pain. He was weaned off mechanical ventilation to a tracheostomy collar but developed respiratory distress and significant oxygen desaturation. This was thought to be secondary to pulmonary edema from congestive heart failure given his previous cardiac history and findings from his exam and diagnostic tests. His electrocardiogram (ECG) was unchanged compared with prior exams, his cardiac enzymes were unremarkable, and his white blood cell count and differential were unremarkable. His sputum was not purulent, and his chest X ray showed diffuse pulmonary infiltrates read by the radiologist and medical consultant as edema rather than a focal infiltrate suggestive of pneumonia. There was no unilateral lower leg edema noted, and he had been on sequential compression devices for deep venous thrombosis prophylaxis. He was treated with albuterol sulfate/ipratropium bromide via nebulizer and diuresed with furosemide at the recommendation of the internal medicine consulting service. His breathing and oxygen saturation improved, but shortly thereafter the patient became tachycardic in the 120–140 bpm range.

The patient's tachycardia was presumed by the primary team and medical consultant to be multifactorial, including an exacerbation of his atrial fibrillation in the setting of diuresis-induced hypovolemia, anemia with a hematocrit of 28, as well as "pain and

anxiety." His ECG showed atrial fibrillation with a rate of 110 bpm but was otherwise not significantly changed. His QTc was 470 msec. Repeated assays of his cardiac troponin I and creatine kinase-MB were within normal limits. He was started on metoprolol, and his amlodipine, benazepril, aspirin, and atorvastatin were continued. His heart rate returned to the 90s to low 100s, and his blood pressure ranged in the previous 24 hours of our consultation from 124/80 to 136/90 mm Hg.

A percutaneous gastrostomy tube had been inserted, as the patient would need to remain on a nothing by mouth (NPO) regimen for several months as a result of the procedure performed. The patient developed abdominal pain with feeding through the tube and was taken back to the operating room for exploratory laparoscopy and replacement after peritoneal extravasation was noted with a contrast study. His medications were switched to intravenous equivalents wherever possible. After the tube was replaced, the patient was able to receive his medications and nutritional support via the gastrostomy tube. He continued to take omeprazole. He had bowel movements daily or every other day, and there was no evidence of impaction.

The patient was febrile on postoperative day 2 to 38.6°C despite receiving clindamycin, which was documented as being prescribed "for prophylaxis" during the surgery. His surgical site appeared unremarkable; his skin was examined and no pressure sores or ulcers were noted. His pulmonary exam and diagnostic tests, as described previously, were not felt to be consistent with pneumonia. Blood, sputum, and urine cultures were sent for analysis. Lumbar puncture was not performed. He was started on ciprofloxacin for what appeared possibly to be a consideration for peritonitis from the extravasation of the gastrostomy tube. On postoperative day 3 the ciprofloxacin was written as being prescribed for a presumed urinary tract infection as 2+ leukocyte esterase was noted on his urinalysis.

From a neurologic and psychiatric perspective, the patient was noted immediately postoperation to be heavily sedated with midazolam and fentanyl. He received additional intermittent hydromorphone hydrochloride for pain. His pain was documented as being "well controlled" and "0/10" in the nursing notes. On postoperative day 2 he was extubated, the midazolam was stopped, and he was started on patient-controlled analgesia with morphine. At that time he was documented to be "alert and oriented × 3" and to be "responding to commands" in the notes. His exam was documented as being "nonfocal." On postoperative day 3, however, he appeared to be "anxious," "depressed," and "tearful." The nurse informed the team that she felt the patient was growing increasingly "upset," possibly as a result of having to deal with several of the previously mentioned medical problems that were developing and were complicating his recovery.

He continued to be "oriented × 3" and "able to respond to commands" and continued to be described as "denies pain." In the evening before our consultation on postoperative day 3, the patient was noted as having become "increasingly tearful, upset, and anxious," gesticulating tremulously with his hands in a manner that appeared to the nursing staff to be insisting that he was "thirsty" and "wanted to drink water." He was receiving intravenous fluids as well as nutrition through his gastrostomy tube. He was NPO and had been instructed repeatedly not to drink. Despite this, he repeatedly reached for water on the nursing tray at the bedside that was intended to flush his gastrostomy tube. He attempted to drink this several times despite repeated instructions to remain NPO before the water was removed from the bedside by the nursing staff. He was tearful and was not consolable. Shortly afterward he was administered lorazepam and hydromorphone intravenously, and he became more sedate.

The nursing staff described feeling frustrated that he was acting in an "anxious and irrational" manner, which might complicate his recovery. He was described as being "restless" overnight and did not sleep well. He would not participate in the tracheostomy care and tube feeding training exercises recommended by his nursing staff the following day, and he appeared withdrawn from communication. On the same day he was later described as appearing "panicky" at times. The team, noting the anxiety that had been mentioned in the preoperative consultation notes, asked the patient if he was feeling depressed, and he seemed to nod "yes" in agreement to this and to questions about whether he felt like life was "not worth living anymore," although he would not elaborate. The psychiatry consult was then requested to better "assess the patient's worsening anxiety and depressed mood," which were "interfering with his ability to participate in tracheostomy and gastrostomy self-care instruction and training."

A specific question asked of the team was "should an antidepressant be initiated?"

Other Significant Findings From Assessment

Being called to assess a patient in a scenario similar to that presented previously is a relatively common and challenging task requested of the consultation-liaison team. This patient appeared to be "upset with his care" and to be suffering from "anxiety and depression" to the primary team. He was known to be in a chronic state of poor health and to have been admitted for treatment of his recurrent laryngeal cancer. He had suffered a series of complications after his procedure. He was unable to speak but appeared tearful and to endorse symptoms of depressed mood, anhedonia, and hopelessness. He appeared at times to be anxious and to have difficulty sleeping. He had thoughts of death and seemed to endorse that his life was not worth living. He initially appeared to be actively refusing reasonable interventions such as remaining NPO to improve his outcome and was then observed to be sullen and withdrawn and to refuse physical therapy and rehabilitation. The psychological stress and existential issues in dealing with severe medical illness and the complications of treatment were high. The medical team felt these symptoms were causing sufficient interference with his recovery and wondered whether an antidepressant should be initiated.

• Before we proceed further, take a moment to briefly write down three to five diagnoses you might consider in this patient. We will return to this later as an exercise in discussing hindsight bias (Bornstein and Emler 2001). Please write these down now.

Diagnosis

Importance of Assigning an Accurate, Reproducible Diagnosis

Clinicians assign diagnoses in their attempts to search for relevant past experiences that might permit them a better inductive forecast of the future and guide their decision making to improve patient outcomes. We see a patient today and consider how well their circumstances and pattern of symptoms might be conceptualized as part of one or more previously operationalized diagnostic syndromes. Our knowledge and experience with these syndromes allow us to make educated guesses as to the options most likely to achieve optimal outcomes for our current patient. The more closely our clinical setting and diagnostic methods are to those of clinicians who have previously conducted carefully controlled clinical research trials on these syndromes, and the closer our patient's experiences are to the syndromes described within these trials, the more confidently we might be able to make inferences from the knowledge gained. Where these things are very dissimilar we are unlikely to be able to confidently forecast future outcomes. A person experiencing a depressive syndrome highly similar to the symptoms reported by persons described as having major depressive disorder in a high-quality randomized controlled trial conducted in a similar setting might benefit from the treatments that this research has found helpful. A person whose pattern of symptoms and signs are more similar, however, to those exhibited by the group whom we conceptualize as having bipolar depression might respond differently to these same interventions. Those whose depressive symptoms are associated with additional symptoms and signs suggesting profound hypothyroidism, or alcohol use or amphetamine withdrawal, or obstructive sleep apnea or hypoactive delirium, and so on, might respond differently still. Thus, where treatment options exist that might result differentially in better or worse outcomes for the patient, it is important to consider carefully the diagnostic process used to assign an accurate, reproducible diagnosis so that we can better guide our decision making (Pauker and Kassirer 1980).

Common Diagnostic Biases Leading to Errors in Assigning Diagnoses

The procedures that we use to assign a diagnosis vary. An approach often taken by clinicians is to assign a diagnosis based on the clinical gestalt suggested by the pattern of symptoms observed in examination of the patient. This has the advantage of being rapid and can be quite accurate when applied by experienced clinicians (Chunilal et al. 2003). There are situations, however, in which this approach is susceptible to a number of biases that can lead to assigning diagnoses very different from those using more optimal examination methods (Bornstein and Emler 2001). One of the most common biases

that can occur resulting in assigning an incorrect diagnosis is failing to account for how common a given diagnosis might be in a given setting. This error of ignoring the pretest probability or base rate of a given diagnosis and being susceptible to being overly influenced by the patient's clinical presentation or test findings is a violation of a mathematical theorem of probabilistic reasoning called *Bayes theorem*. This bias, sometimes referred to as the *representativeness bias*, is a well-documented source of diagnostic error (Bornstein and Emler 2001; Casscells et al. 1978; Ghosh et al. 2004; Kahneman et al. 1982; Lyman and Balducci 1993, 1994). In medicine we use the phrase "when you hear hoofbeats, think of horses, not zebras" as a reminder to ourselves not to forget to incorporate the base rates of a given diagnosis in a particular setting. Applying this clinically, however, is challenging even with training (Steurer et al. 2002). Knowing how to combine estimates of the pretest probability of a given diagnosis in a particular setting together with the clinical features that distinguish this diagnosis from others in the differential is part of the art of practicing evidence-based medicine that we will illustrate with this case.

Differential Diagnosis

When Should a Differential Diagnosis and Further Diagnostic Testing Be Considered?

In psychiatry, our clinical trials most often use the procedures suggested by DSM-IV-TR (American Psychiatric Association 2000) to assign diagnoses. The patient in this case certainly did appear to have many symptoms similar to the pattern described as being present in a DSM major depressive episode.

This diagnosis appeared to be a good "snap fit" to the primary team. Would there be any point in generating a differential diagnosis? Major depressive disorder was certainly a possibility. The pattern was clearly not an exact match, but similar features certainly were present.

A second bias often affecting clinical decision making is called the *confirmation bias*. This occurs when we selectively gather and interpret evidence that confirms a diagnosis and ignore evidence that might disconfirm it. Another bias that often occurs affecting the validity of a diagnosis is called the *availability bias*. This bias occurs when we overestimate the probability of a diagnosis when instances of that diagnosis are relatively easy to recall. Generating a differential diagnosis is a method that we use to combat these biases.

In this case, many diagnoses possibly were consistent with this patient's symptoms and circumstances. Numerous serious medical conditions were present, and multiple medications were being prescribed to this patient in the intensive care unit (ICU) in which the delirium syndrome is known to be quite prevalent (Ely et al. 2004). Delirium can present with profound anxiety and depressive symptoms and is often overlooked (Farrell and Ganzini 1995). Treatments for delirium would involve actively seeking and correcting the medical causes of these symptoms, and this would differ markedly from an approach recommended for major depressive disorder. Persons with preexisting dementia, anxiety disorders, mood disorders, or multiple other psychiatric conditions might have increased difficulty coping with the extreme stresses present in this case. We knew nothing of this patient's prior history and whether the symptoms observed were acute or chronic. Substance- and medication-related effects as well as their withdrawal could cause or exacerbate the symptoms exhibited. Any number or combination of factors potentially could be contributing to the symptoms that our patient was exhibiting.

A *diagnostic testing threshold* (Pauker and Kassirer 1980) had therefore been reached. The treatment recommendations varied widely for the conditions that could be present, and we were not convinced that major depressive disorder was the best or the only diagnosis consistent with this patient's symptoms. We had not yet crossed a *test–treatment threshold* in which we could feel confident in our making treating recommendations. This prompted us to expand the differential diagnosis beyond the diagnosis of a major depressive episode and to conduct further diagnostic assessment. In general, as per Pauker (Pauker and Kassirer 1980), treatment should be withheld if the probability of disease is smaller than the diagnostic testing threshold, and treatment should be given without further testing if the probability of disease is greater than the test–treatment threshold. Further diagnostic testing should be considered (with treatment depending on the test outcome) only if the probability of disease is between the two thresholds.

How might a differential diagnosis be generated? A final bias that must be considered in diagnostic decision making is called *regret bias*. This occurs when

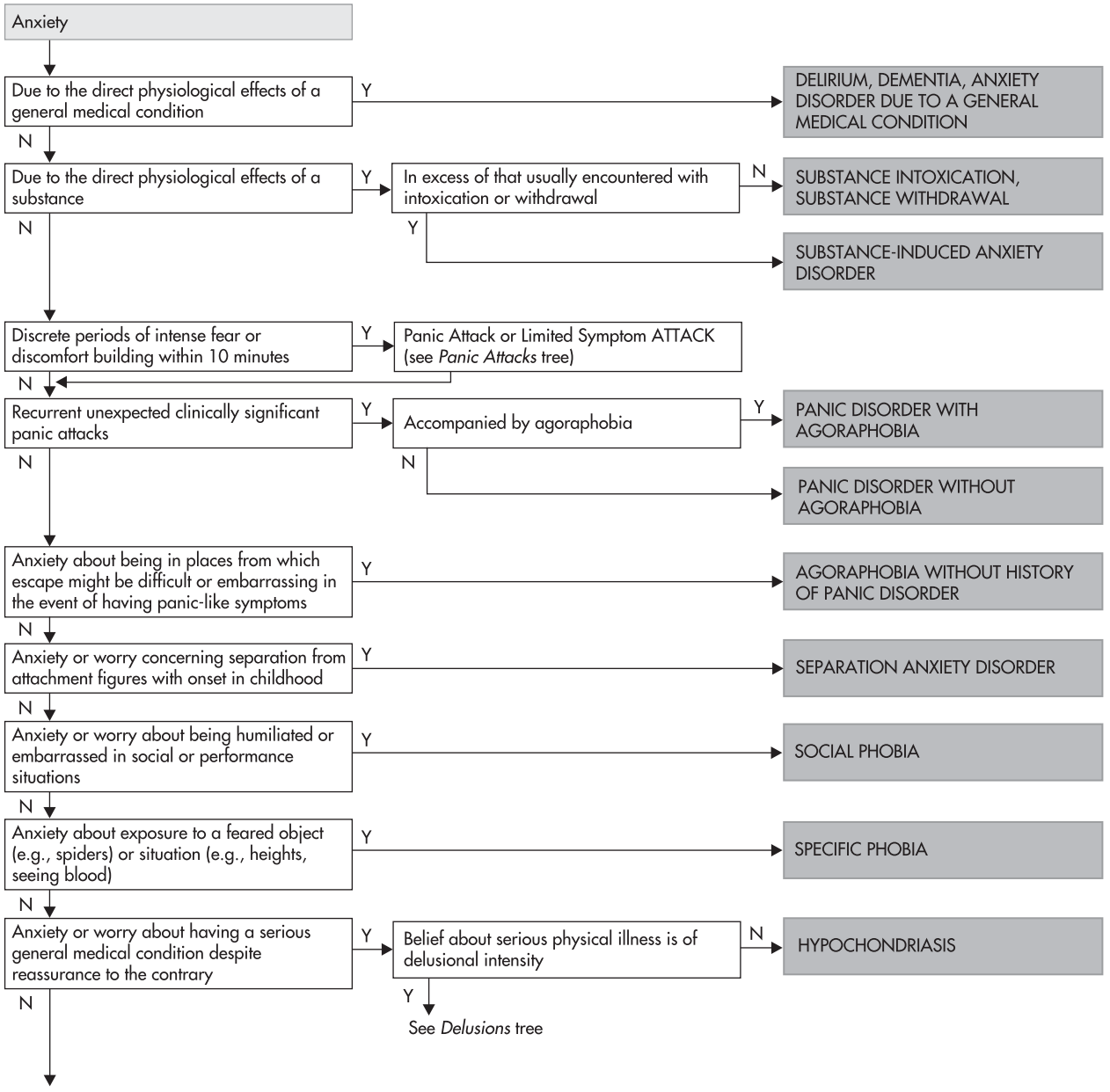


FIGURE 33–1. DSM-IV-TR decision tree for differential diagnosis of anxiety.

Source. First MB, Frances A, Pincus HA: *DSM-IV-TR Handbook of Differential Diagnosis*. Washington, DC, American Psychiatric Publishing, 2002, pp. 21–23. Copyright ©2002 American Psychiatric Publishing, Inc. Used with permission.

we overestimate the probability of a diagnosis with a severe possible outcome because of anticipated regret if such a diagnosis were missed (Bornstein and Emler 2001). We can consider an approach that borrows elements of several potential strategies that a reasonable clinician might use to form a differential diagnosis (Richardson et al. 1999) that minimizes the risks of the various biases introduced previously. We might consider a broad differential diagnosis (possibilistic approach) that weights more heavily those causes that cannot be missed (prognostic approach) as well as those that are highly likely in that practice setting (probabilistic approach) that can be treated resulting in better patient outcomes (pragmatic approach).

The *DSM-IV-TR Handbook of Differential Diagnosis* (First et al. 2002) is a resource that we might consider consulting to help us form and focus our initial differential diagnosis. It includes decision trees (see Figure 33–1) that help the psychiatric clinician to use

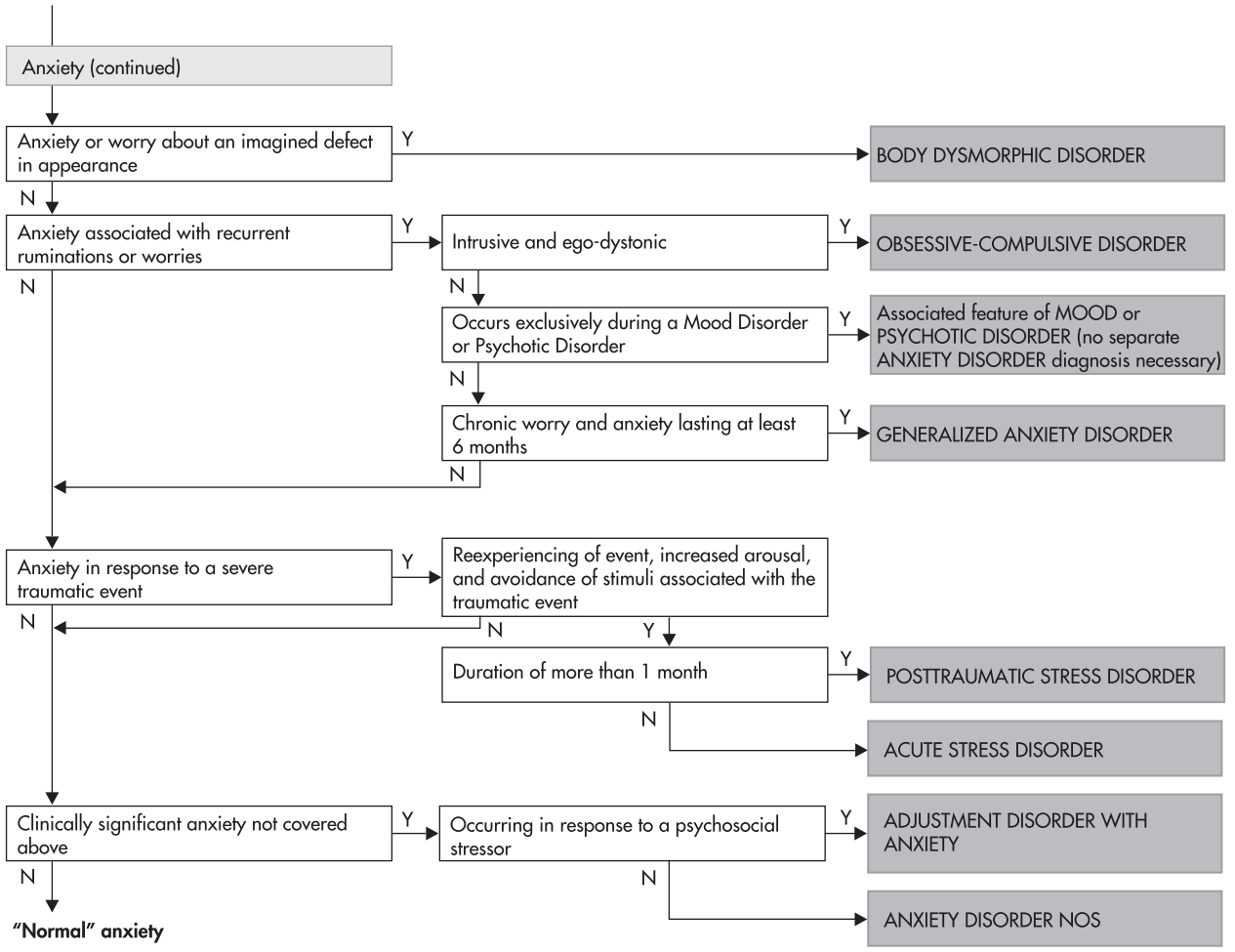


FIGURE 33–1. DSM-IV-TR decision tree for differential diagnosis of anxiety (*continued*).

procedures to assign diagnoses consistent with those of the formal procedures described in the text of the DSM used in randomized controlled trials. It is organized somewhat consistently with the pragmatic strategy described previously, weighting those conditions such as disorders commonly caused by medical or substance-related etiologies as diagnoses to examine prior to assigning other diagnoses because these syndromes might have important treatment implications. The clinician must use this tool carefully because many nuanced factors must be considered together with this resource, as we will see.

Delirium, dementia, and anxiety secondary to a general medical condition or pharmacological substance intoxication/withdrawal are the diagnoses that appear at the top of the list from the *DSM-IV-TR Handbook of Differential Diagnosis*. Delirium is an extremely prevalent state of acutely altered mental status that occurs in patients who are seriously ill. Occurrence rates of delirium in hospitalized medical patients range from 11% to 42% of all admitted patients (Siddiqi et al. 2006) on general medical units. Estimates are much higher for those in the ICU, where our patient consultation occurred. For example, a well-designed prospective study using validated methods for detecting delirium found that 81.7% of mechanically ventilated patients in the ICU were delirious (Ely et al. 2004). Despite the high prevalence rates, clinicians are often led astray by the clinical presentation of delirium. Delirious patients often endorse anxious and depressive symptoms, such as low mood (60%), worthlessness (68%), and frequent thoughts of death (52%) (Farrell and Ganzini 1995), as part of the delirium syndrome. Patients who are alert, easily aroused, able to make eye contact, and able to follow commands are often presumed by clinical staff to have normal cognitive functioning, but 40% of these patients have

been found to be delirious when examined by more rigorous examination methods (Ely et al. 2001a). Acute changes in mental status in elderly patients in medical settings frequently are misattributed to primary psychological causes such as anxiety or depression (Boland et al. 1996; Farrell and Ganzini 1995; Nicholas and Lindsey 1995; Swigart et al. 2008), if they are diagnosed at all (Gustafson et al. 1991; Inouye 1998; Inouye et al. 2001). Delirium is an important risk factor for increased mortality in both general medical (Siddiqi et al. 2006) and ICU settings (Ely et al. 2004) and has very different treatment implications than the other diagnoses considered in the differential. Delirium can be a marker of serious substance withdrawal syndromes and other acutely unfolding medical conditions and thus became a very important diagnosis for us to consider in our differential diagnosis.

DSM-IV-TR diagnostic criteria for delirium appear in Table 33–1.

TABLE 33–1. DSM-IV-TR diagnostic criteria for delirium due to multiple etiologies

A.	Disturbance of consciousness (i.e., reduced clarity of awareness of the environment) with reduced ability to focus, sustain, or shift attention.
B.	A change in cognition (such as memory deficit, disorientation, language disturbance) or the development of a perceptual disturbance that is not better accounted for by a preexisting, established, or evolving dementia.
C.	The disturbance develops over a short period of time (usually hours to days) and tends to fluctuate during the course of the day.
D.	There is evidence from the history, physical examination, or laboratory findings that the delirium has more than one etiology (e.g., more than one etiological general medical condition, a general medical condition plus substance intoxication or medication side effect).

Source: American Psychiatric Association 2000, p. 147. Used with permission.

Narrowing the Differential

The patient at the time of our examination was unfortunately unable to speak to us as a result of his recent laryngectomy. There was little information available in the chart record about any past psychiatric history, cognitive impairment, dementia, or substance use, and no family members were present at the bedside. We did not know if the patient lived independently or in an assisted living facility or if he had caregivers at home. This is a common scenario that can make assessment more challenging. Often in these cases attempting to collect information from collateral historians and conducting a thorough physical and mental status exam are the only sources of information available.

DSM-IV-TR diagnostic criteria for delirium require that a preexisting dementia be excluded as being better able to account for the disturbances observed. Because delirium was the most critical diagnosis in our differential, evaluating for preexisting dementia thus became important. Treatment recommendations for agitation in dementia could be considerably different from those recommended for delirium, which focus much more intensely on identifying and primarily correcting underlying physiologic processes and secondarily on pharmacotherapies to help manage agitation. If the patient were delirious from alcohol or benzodiazepine withdrawal, for example, we would recommend a substantially different treatment than we would for a patient agitated from Alzheimer's disease. Thus a diagnostic testing threshold had been reached to tease these apart. The absence of dementia would make our search for underlying reversible physiologic causes much more important. The presence of dementia of course would not exclude a delirium upon dementia and would result in a much more careful and detailed examination before we felt comfortable in excluding this possibility. Unfortunately, detection of dementia in clinical settings is difficult (Ardern et al. 1993; Boustani et al. 2003; Chodosh et al. 2004; Holsinger et al. 2007; Olafsdottir et al. 2000; Pisani et al. 2003b; Roca et al. 1984; Valcour et al. 2000). The absence of mention of dementia in the medical record was unlikely to exclude this as a diagnostic concern.

Distinguishing Dementia and Delirium From Other Diagnoses in the Differential

An Introduction to Likelihood Ratios, Pretest Probability, and Evidence-Based Diagnosis

In an ideal scenario, the clinical judgment of the experienced physician is informed by the pretest probabilities of some of the more important potential causes in the differential diagnosis that might occur in that specific practice setting (Richardson et al.

1999, 2000a, 2000b, 2003). Diagnostic procedures with operating characteristics derived from high-quality diagnostic studies conducted in similar settings (Bossuyt et al. 2003b) might be used to assist the clinician in sorting through the differential. *Likelihood ratios* are the operating characteristics of various clinical findings (including items in the history, physical exam, or diagnostic tests) that we use in combination with pretest probabilities to estimate the *posttest probability* of a given condition being present (Deeks and Altman 2004; Sackett 1992).

Rarely are these findings capable of ruling in or ruling out a diagnosis by themselves. Likelihood ratios are weighting factors that tell us how strongly the presence or absence of a given finding is in influencing our estimate of the probability of a given disorder being present. Understanding these operating characteristics is important in assisting us in overcoming representativeness bias. Several studies suggest that we often grossly overestimate the findings of the clinical presentation and diagnostic tests in our decision making (Ghosh et al. 2004; Lyman and Balducci 1993, 1994; Rothman and Kiviniemi 1999; Sackett 1992; Steurer et al. 2002; Streiner 2003). A likelihood ratio greater than 1 indicates that a given finding is associated with the presence of the disease, whereas a likelihood ratio less than 1 indicates that a given finding is associated with the absence of disease. The further likelihood ratios are from 1, the stronger the evidence for the presence or absence of disease. Likelihood ratios above 10 and below 0.1 are considered to provide strong evidence to rule in or rule out diagnoses, respectively, in most circumstances; however, the pretest probability must always be considered. Diagnostic procedures for which likelihood ratios are available include assessment of the symptoms endorsed in the patient's history, mental status exam, or conducting structured bedside assessment inventories such as the Folstein Mini-Mental State Examination (McGee 2007). Likelihood ratios are easily calculated from the same information used to calculate sensitivity and specificity. The Rational Clinical Exam series in *JAMA* (Sackett 1992) has dozens of articles dating back over a decade presenting likelihood ratios and pretest probabilities for many of the most common medical problems that we encounter clinically, although the quality of many of the available diagnostic studies remains an ongoing source of concern (*Bandolier* 2002; Richardson 2007).

Does the Patient Have Dementia?

We knew from a well-designed prospective study using validated methods for detecting delirium that more than 80% of mechanically ventilated patients in the ICU were delirious (Ely et al. 2004) The prevalence of dementia is not well studied in inpatient medical settings. Our previous exploration of the medical literature had yielded only one study that we felt provided a reasonable estimate of the prevalence rate of preexisting (nondelirious) cognitive impairment in the ICU (Pisani et al. 2003a). This study estimated the pretest probability to be 42%. If we make the assumption for the time being that this estimate was a reasonable one, we then are faced with the task of identifying and implementing further diagnostic procedures with known operating characteristics by which to modify our pretest probability estimates.

Examining a nonverbal patient in the ICU under medical duress for dementia is not an easy task. We hoped to find diagnostic procedures using collateral historians as informants to assist us in our task. Because this situation arises so often and has important diagnostic implications, we previously have conducted a focused search of the literature to find an answer to the PICO question below (see Chapter 3 of this volume and Richardson et al. 1995).

Ask

Patient/population: Medically ill elderly hospitalized patients suspected of having dementia.

Intervention: Brief, structured, informant-based interview.

Comparison: Much more intensive "gold standard" reference examination integrating information from collateral historians, prehospital cognitive performance–based measures, and expert clinical evaluation.

Outcome: Accuracy—likelihood ratios (positive and negative) of the brief measure compared against the gold standard. Precision—measures of interrater reliability such as kappa or intraclass correlations.

Acquire

There are unfortunately no organized repositories of high-quality preappraised systematic reviews of the medical literature pertaining to diagnosis available for

rapid clinical queries at the time of this writing. The Rational Clinical Examination series in *JAMA* contains a limited review of diagnostic methods for various clinical problems but does not have a rapidly accessible search interface. The Cochrane database is, at the time of this writing, still in the process of developing a series of systematic reviews pertaining to diagnostic studies. As a result we do not expect to conduct such a search at the bedside to find a rapid answer to our question. We instead do monthly searches and appraisals of articles pertaining to diagnosis as part of a journal club in which we present critically appraised topics (CATS) relevant to our clinical work, then save these to a Web site for easy access during future episodes of patient care (Mascola 2008; Sackett and Straus 1998). Previously we had searched the literature for high-quality systematic reviews addressing the topic in the PICO question described previously using the PubMed Clinical Queries filters and were unable to find any reviews focused on assessing hospitalized patients for dementia. We had found previously two systematic reviews of studies conducted in primary care settings that referenced a limited number of informant questionnaires for assessing for dementia in outpatient settings (Boustani et al. 2003; Holsinger et al. 2007).

The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) was an instrument that appeared in these reviews to be brief, requiring only minutes to perform, that collects information from collateral historians. It appeared to have been validated against several more resource-intensive, performance-based criterion standards for dementia (Jorm 2004) in outpatient settings in which such validation would be more easily performed and also to have been used in prognosis studies in inpatients with delirium (McCusker et al. 2002). It had favorable operating characteristics in these settings relative to these gold standard examinations. We thus focused our search for trials in which this measure might have been validated in medically ill inpatients and found one study in which the measure had been used in a teaching hospital general medical unit (Harwood et al. 1997) and another in which it was used to assess patients in the ICU (Pisani et al. 2003a). We had selected these articles in previous months to appraise critically for validity and applicability in our practice setting because we felt they might affect our practices, and we had CATS and the measures accessible on our internal Web site.

Appraise

We often use the STARD (Standards for the Reporting of Diagnostic Accuracy Studies) checklist for the reporting of studies of diagnostic accuracy (Bossuyt et al. 2003a, 2003b) to appraise the diagnostic studies that we find for quality in writing up our CATS. The STARD checklist is meant for editors of medical journals to use in appraising studies of diagnostic accuracy for completeness of reporting and for common sources of bias that can invalidate the findings. It works nicely for our purposes. There are other instruments available to use to appraise studies of diagnostic accuracy for quality (Richardson et al. 1999; Whiting et al. 2006) that you might consider in your setting.

- The reader is encouraged to stop here and obtain the Harwood study (Harwood et al. 1997), if fortunate enough to have access to the full text. You can then use a copy of the STARD checklist as an exercise in appraising a diagnostic study. (You can obtain the checklist at www.stard-statement.org and the guide for its use from Bossuyt et al. 2003a.) You might consider taking 30 minutes or so to read through the study while completing the checklist using the how-to guide. You can then compare your appraisal of the quality of the Harwood study with our appraisal and determine whether you might consider using the IQCODE in the assessment of the patients you see in your setting. Continue reading after you have finished your appraisal.

The Harwood study estimated the IQCODE had an optimal post hoc sensitivity of 100% and specificity of 86% using a cutoff of >3.44 as positive for dementia. The reader can convert these sensitivities and specificities into likelihood ratios by using the following formulas (McGee 2007):

$$\text{LR+} = \text{sensitivity}/(1-\text{specificity}) \text{ and}$$
$$\text{LR}- = (1-\text{sensitivity})/\text{specificity}$$

These formulas are available on many of the online EBM calculator sites. Crunching the numbers yields a positive likelihood ratio (LR+) of 7.1 and a negative likelihood ratio (LR–) of 0.1 against a gold standard diagnosis made by detailed informant history together with clinical evaluation by an examiner using DSM criteria. This means that conducting a brief informant interview and obtaining a score above 3.44 on this instrument would modestly increase the probability of dementia being present, whereas scores

<3.44 on this instrument would be likely to greatly reduce the probability of a diagnosis of dementia in a patient being examined. These are impressive operating characteristics for a brief test conducted in a challenging clinical setting. We thus wanted to appraise the Harwood study to see if the estimates of these operating characteristics were likely to be valid and generalizable to our clinical setting.

In our critical appraisal we felt that this study fell short of optimal STARD criteria and included appreciable risk for bias, which could result in the test characteristics appearing much better than they might actually be at discriminating demented from nondemented patients. We present the most important concerns we had from our appraisal using STARD below. Our comments correspond to the item numbers on that checklist. We focus, because of space constraints, on those items that, to our read, conferred a significant risk for bias.

Items 8, 9, and 10 from the STARD checklist (Table 33–2) regarding the reference standard diagnosis were of concern to us. The final process by which the reference standard diagnosis (DSM diagnosis of dementia or clinical diagnosis of other cognitive impairment) was assigned in this study was difficult to determine. The qualifications, expertise, and training of the person executing and reading the index (IQCODE) and reference standard tests were not specified in the body of the manuscript. The qualifications of the author cited as doing the assessment could be searched on the Internet, and he appeared to be an accomplished authority in his field. One challenge, despite the authority of the author, was to understand the process that he used to come to the gold standard diagnosis so that we could understand how accurate the estimate of the test's performance in agreement with this diagnosis was likely to be. STARD reminds us that this process should be clearly specified and reproducible so that the results of the research are replicable and the procedures used to reach a gold standard diagnosis do not introduce bias into the estimate of test accuracy. The procedures used in the study listed the tests performed for the reference standard but included sources of history and procedures for gathering information that were poorly specified and would be difficult to reproduce with the potential for introducing significant bias.

For example, in appraising the study for STARD item 11 we felt the procedures used introduced the possibility of a form of bias referred to as *clinical review bias*, in which interpretations of diagnostic procedures become more accurate by providing additional clinical information to interpreters. The methods section of the study describes that a single person, the author mentioned previously, performed the Confusion Assessment Method (Inouye et al. 1990) to detect delirium, the Abbreviated Mental Test (Hodkinson 1972), and the 16-item version of the IQCODE as well as "further standard psychiatric assessment…[which] included further cognitive assessment, obtaining a history from the informant and/or nursing staff and scrutinizing clinical notes." The additional clinical information provided by each test was likely to influence the findings of the other tests performed.

The author stated that "all patients with some evidence of abnormal cognitive function—either a positive screening test or a note indicating abnormal cognitive status in the clinical records—were assigned a DSM-III-R diagnosis, or if not meeting DSM criteria, a clinical diagnosis of their cognitive impairment." It appeared from this description that results of the IQCODE influenced whether the gold standard diagnostic procedures were performed and whether a DSM-III-R or clinical diagnosis was assigned, a source of bias referred to as *verification bias* (item 16), that could increase the estimate of diagnostic accuracy (Bossuyt et al. 2003b). The person assigning the criterion standard diagnosis was not blinded (item 11) to the findings of the index test, which were included in the gold standard diagnosis reached. The same person was assigning both diagnoses. This results in both *test review bias* and *diagnostic review bias*, both of which are likely to inflate the estimates of diagnostic accuracy. No confidence intervals about the point estimates for sensitivity and specificity were provided (item 21), nor was a cross-tabulation of the results of the index tests by the results of the reference standard given (item 19). The operating characteristics provided used optimal post hoc estimates, which were unlikely to be replicated. However, cutoff points suggested by the authors as the maximally discriminative values obtained from previous outpatient validation studies of the test with more rigorous designs had similar accuracy (test score cutoff >3.38 had a sensitivity of 100%, specificity 84%). There were no estimates of the variability of diagnostic accuracy or reproducibility provided.

Overall we felt that further study of this test with respect to a gold standard test using a more rigor-

TABLE 33–2. STARD checklist of studies of diagnostic accuracy

Section and topic	Item #		On page
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading "sensitivity and specificity").	
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	
METHODS		Describe:	
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	
Test methods	7	The reference standard and its rationale.	
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	
	13	Methods for calculating test reproducibility, if done.	
RESULTS		Report:	
Participants	14	When study was performed, including beginning and end dates of recruitment.	
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	
	20	Any adverse events from performing the index tests or the reference standard.	
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	
	22	How indeterminate results, missing data and outliers of the index tests were handled.	
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	
	24	Estimates of test reproducibility, if done.	
DISCUSSION	25	Discuss the clinical applicability of the study findings.	

Note. MeSH=Medical Subject Heading; STARD=Standards for Reporting of Diagnostic Accuracy.

Source. Bossuyt et al. 2003a. Available at: http://www.stard-statement.org.

ously controlled design would be likely to result in a change in the estimate of diagnostic accuracy and that the estimate provided in this study was likely to be inflated. The study performed in the ICU to estimate the prevalence of preexisting cognitive impairment (Pisani et al. 2003a) indicated that there was substantial agreement between the IQCODE and the Modified Blessed Dementia Rating Scale—another informant interview validated in outpatient settings—with a kappa of 0.69. This study was limited by the difficulty in establishing a gold standard diagnosis against which to estimate the operating characteristics of the test.

Apply

These studies appeared to be the only recent studies looking at methods to determine whether a patient might have a preexisting, nondelirious form of cognitive impairment, such as dementia, in the setting in which we found our patient. For our purposes we decided that despite the limitations, the IQCODE remained the best-studied instrument available for our purpose and that the point estimates derived from this limited study would represent our best estimate as to this test's accuracy. We felt that the accuracy estimate was likely to be biased in favor of greater accuracy than was warranted, and thus we made a mental note of this in our interpretation.

We thus contacted the patient's family to obtain their impression of the patient's mental state at baseline prior to admission and to administer the IQCODE. The patient had been married for 53 years. His wife was in her early 70s and was a retired administrative assistant for the local technology company where the patient had worked for many years. She had been with the patient daily prior to the hospitalization. He also had a daughter in her mid-40s who was a schoolteacher in the local area. Both appeared to be good historians and had spent a minimum of 4 hours per week with the patient for at least 10 years as was suggested by the IQCODE procedure.

They were contacted for interview over the telephone. They informed us that they had just the left the hospital and were going to get something to eat after being at the bedside for many hours that day as they had each day since admission. They were quite concerned about the patient's mental state. They had never known the patient to be depressed or anxious previously, and he had never before seen a psychiatrist for any reason. His wife felt that he was in good spirits prior to the admission and he seemed to be dealing fairly well with the knowledge that his cancer had returned. He had expressed hopefulness and confidence in the surgical approach recommended by his doctors, although he was anxious as to whether his heart would tolerate the procedure. After receiving the diagnosis he seemed to still be able to enjoy visiting with friends and family as he had before. He still paid the bills for the household and was quite active. He had been productively employed as an engineer at a local technology company until about 8 years before he was admitted. He had slowed down somewhat after his heart surgery in 1991, but he still managed to enjoy his activities and to manage his affairs without difficulty. Neither she nor her daughter had noticed any significant decline in his memory or ability to take care of things around the house in the years preceding his current admission. They described him as being "as sharp as ever" before his hospitalization. We went through each item on the 16-item IQCODE together, comparing his cognitive abilities 1 month prior to his admission to his performance 10 years before as was done in Pisani et al. (2003a). The final score was 3.1, which was a negative result using the cut point of 3.44 suggested by the authors.

Remember that the Bayes theorem tells us that the findings of the test must be considered together with the pretest probability of the condition before a diagnosis is assigned. Remember representativeness bias. This is the clinical error that occurs when we assume that test results rule in and rule out diagnoses without considering the probability of the diagnosis and the fact that each test has false positives and negatives and is not perfect. We use information from the pretest probability of dementia together with the negative findings of this test to estimate our posttest probability of dementia.

We assumed a pretest probability of dementia for hospitalized patients in the ICU of 42% from the best estimate we had available (Pisani et al. 2003a). We had an optimistic estimate of the operating characteristics of the IQCODE from the Harwood et al. (1997) study, which provided a negative likelihood ratio of 0.1. We combined these pieces of information to estimate the posttest probability of dementia using an adaptation of Fagan's nomogram (Fagan 1975; Glasziou 2001). See the introductory chapters from McGee (2007) or the article by Sackett (1992) for more on how to do this if this is confusing to you. This approach resulted in an estimate of the posttest

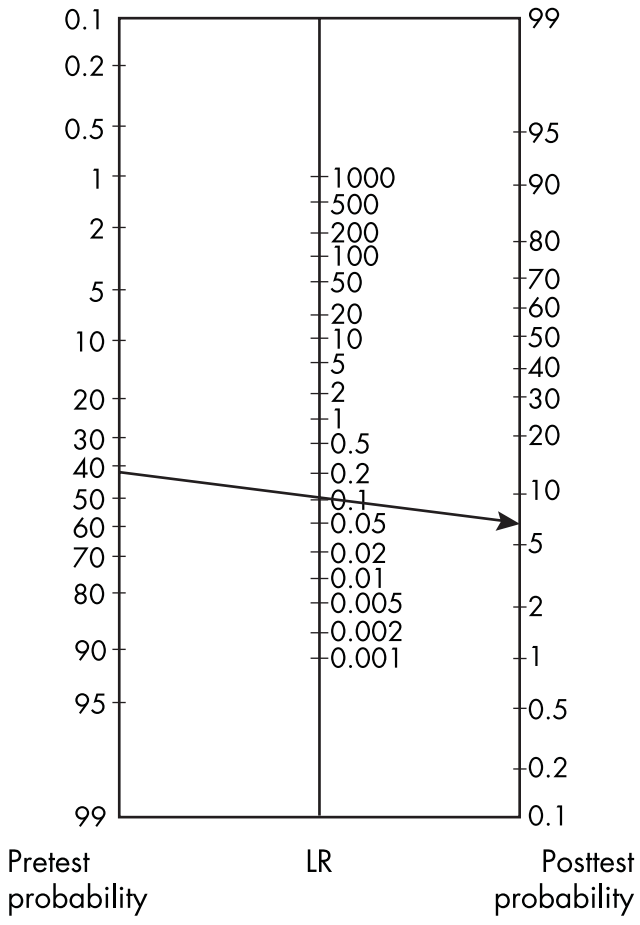


FIGURE 33–2. Likelihood ratio (LR) nomogram for posttest probability of dementia.

probability of dementia as being fairly unlikely, in the range of 5%–10%, as plotted in Figure 33–2. (An online automated version of Fagan's nomogram can be found at www.cebm.net/index.aspx?o=1161.)

Because we had previously appraised this literature and had this information easily available to us, our rough estimate of the probability of dementia took less than 10 minutes to acquire, including our conversation with the patient's wife, which took the bulk of the time, and the previous calculations using the nomogram, which took approximately 10 seconds. While the precision of the quantitative estimate of the probability of dementia was somewhat uncertain given the biases likely to be present in the validation of the test, we were confident in our clinical judgment at this point that the history pointed to a much more acute change in the patient's mental status that would be much more consistent with delirium than with dementia. The results certainly had clinical face validity, and the focused questions in the assessment included in the IQCODE were likely to have prompted us to consider more carefully symptoms that we might not otherwise have assessed for without a structured questionnaire. We decided to continue to investigate for the possibility of delirium being present in our patient.

Does the Patient Have Delirium?

Given the frequency with which we encounter delirium in our practice setting we also hoped to have a soundly validated procedure with known operating characteristics that we might use together with estimates of pretest probability to provide an accurate assessment of the probability of delirium being present in our patient. We had thus formed and sought an answer to the following PICO question.

Ask

Patient/population: Patients in intensive care unit settings.

Intervention: Brief, bedside diagnostic methods for assessing for delirium.

Comparison: More extensive expert clinical consensus diagnosis using DSM criterion.

Outcome: Outcome measures as before—likelihood ratios positive and negative, estimates of interrater reliability such as kappa.

Acquire

In our previous searches of the medical literature we found a handful of studies addressing this question. One method developed for the ICU for assessing mechanically intubated, nonverbal patients is the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) (Ely et al. 2001a, 2001b).

• We encourage the reader to stop here and obtain a copy of the article in *JAMA* for the validation of this study (Ely et al. 2001a), if fortunate enough to have full-text access. Please also print out the STARD checklist and critically appraise this and see if you agree with our summary below. No peeking!

Appraise

The validation study for this instrument in our opinion was quite strong with respect to the STARD criteria (Ely et al. 2001a). By going through the checklist we detected very little risk for bias in the methods used. The validation study did exclude assessment of patients with a previous history of psychosis and neurologic disease (limited challenge bias), of which we made note. The study did not provide a cross-tabulation of the results of the index tests by the results of the reference standard, reporting the summary operating characteristics only. Overall, however, the study addressed each of the major threats to validity for studies of diagnostic accuracy. The use of the CAM-ICU appeared to be significantly more rapid (a mean of 2 minutes, with a standard deviation of 1 minute) than the gold standard assessment (30–45 minutes) using usual clinical methods of arriving at a DSM-IV-TR diagnosis. Overall our critical appraisal by the STARD criteria suggested future studies of this instrument would be unlikely to result in a significant change in the operating characteristics reported. The positive likelihood ratio for this test was quite strong, with point estimates for each assessor at 50 (95% confidence interval [CI] 20–77) and >100 (95% CI 21 to >100). The negative likelihood ratio was not reported. Using the sensitivity and specificity data we estimated point estimates for the LR– to be 0 and 0.07, respectively. It had excellent interrater reliability, with a kappa of 0.96 (95% CI 0.92–0.99). We were more confident in the estimates provided by the validation report of this test than in those provided by the IQCODE as a result of the stronger design and reporting of the study.

Apply

Training information (www.icudelirium.org) was available that led our team to feel confident in our ability to use the CAM-ICU in our setting. We have thus adopted this instrument into our routine assessment of mechanically ventilated patients in the ICU for delirium, but do note that this test has not been validated in patients with preexisting cognitive or neurologic impairment and that its operating characteristics in these populations are likely to be significantly less strong.

The patient's wife and daughter described having visited with him daily since the operation. They noted that shortly after the surgery they had noticed a marked change in his emotions and behavior that were very troubling to them. At times the patient seemed to be frightened, suspicious, and confused. At other times he seemed to be very tearful. He was often inattentive and would nod off to sleep. He was "trembling" and "sweaty." They were quite concerned to see him crying and nodding his head when the doctors asked him if he was depressed and hopeless, because he had always been a very optimistic man. When they tried to console him and talk to him about how he was feeling they found that he seemed to be confused and agitated, which appeared very out of character for him. They had never previously noticed that he had any problems with his memory, but he seemed to be unable to recall important information or to be confused, although it was never easy to know because he was unable to speak.

The nursing staff from the night shift prior to our consultation had noted that the patient attempted to drink water at the bedside despite being NPO. The patient's wife and daughter described this behavior as being very out of character for him. The vigorousness of his attempts to obtain water and get out of bed had appeared to prompt his receiving lorazepam

and hydrocodone. At other times, according to the records provided by the physical therapist, he appeared sullen, apathetic, and withdrawn.

In performing the CAM-ICU (Figure 33–3) we noted that the Richmond Agitation-Sedation Scale (RASS) score ranged between +3 on the evening prior to our examining the patient when he attempted to get out of bed and to obtain water to –1 at the time we initially came to see him in which he was drowsy but able to open his eyes to verbal stimulation and keep them open for more than 10 seconds. This, together with the history obtained by the patient's wife and daughter, satisfied the acute onset or fluctuating course criterion for us. He exhibited marked difficulty with attention and had markedly disorganized thinking on these items of the CAM-ICU, answering inappropriately to questions such as "Are there fish in the sea?" "Does a stone float on water?" "Can you use a hammer to pound a nail?" He also exhibited difficulty in following the command. Thus we felt the CAM-ICU overall was positive and we were then able to proceed to estimate a quantitative probability of delirium being present in this patient.

Pretest probability estimates of delirium in studies using the CAM-ICU excluding patients with neurologic illness and psychosis range between 22% and 87%, with most of the estimates being closer to the high end of the range (83%, 87%, 47%, 22%) (Devlin et al. 2007). Assuming even the most conservative pretest probability estimate of 22% and the more conservative of the two-point estimates of the likelihood ratio positive of 50, the posttest probability for delirium is >90% (Figure 33–4).

We thus felt that it was extremely likely that our patient was in fact delirious and that the cognitive, emotional, and behavioral symptoms should not at this time be assumed to be secondary to depression as the primary team had been considering prior to our consultation. We felt that the depressive symptoms were more likely to be those that are often displayed in conjunction with the delirium syndrome that had been previously reported (Boland et al. 1996; Farrell and Ganzini 1995; Nicholas and Lindsey 1995; Swigart et al. 2008).

What is the Differential Diagnosis of the Delirium Syndrome?

A recent guideline on delirium written consistently with the conventions established by AGREE (Appraisal of Guidelines Research and Evaluation; AGREE Collaboration 2003) was consulted for guidance (British Geriatrics Society 2006). Our assessment process was not yet complete because the most important action for the management of delirium is the identification and treatment of potential underlying reversible causes (British Geriatrics Society 2006). The differential diagnosis of the many suspected causes of delirium is quite lengthy and poorly studied, and a detailed discussion is beyond the scope of the space we have. Many risk factors have been described for the development of delirium (mostly from non-ICU cohorts), although none has been proved with a strong association or definite causality (Devlin et al. 2007; Kraemer et al. 1997). The pretest probabilities of these risk factors are poorly defined, making workup challenging.

The underlying cause of delirium is often considered to be multifactorial. Common contributory medical causes of delirium include (British Geriatrics Society 2006):

- Infection (e.g., pneumonia, urinary tract infection)
- Cardiac illness (e.g., myocardial infarction, heart failure)
- Respiratory disorder (e.g., pulmonary embolus, hypoxia)
- Electrolyte imbalance (e.g., dehydration, renal failure, hyponatremia)
- Endocrine and metabolic disorder (e.g., cachexia, thiamine deficiency, thyroid dysfunction)
- Drugs (particularly those with anticholinergic side effects [e.g., tricyclic antidepressants, antiparkinsonian drugs, opiates, analgesics, steroids])
- Drug (especially benzodiazepine) and alcohol withdrawal
- Urinary retention
- Fecal impaction
- Severe pain
- Neurological problem (e.g., stroke, subdural hematoma, epilepsy, encephalitis)
- Multiple contributing causes

In Mr. C.K.'s case, several of these possible risk factors were known to be present. We anticipated that identification of a single culprit that caused his delirium might be unlikely, but we felt it could be important to see if potential causes could be identified, for which treatment had not been initiated, that if treated might reduce his risk for harm.

CAM-ICU Worksheet

Feature 1: Acute Onset or Fluctuating Course Positive if you answer 'yes' to either 1A or 1B.	**Positive**	**Negative**
1A: Is the pt different than his/her baseline mental status? Or **1B:** Has the patient had any fluctuation in mental status in the past 24 hours as evidenced by fluctuation on a sedation scale (e.g. RASS), GCS, or previous delirium assessment?	**Yes**	**No**
Feature 2: Inattention Positive if either score for 2A or 2B is less than 8. Attempt the ASE Letters first. If pt is able to perform this test and the score is clear, record this score and move to Feature 3. If pt is unable to perform this test or the score is unclear, then perform the ASE Pictures. If you perform both tests, use the ASE Pictures' results to score the Feature.	**Positive**	**Negative**
2A: ASE Letters: record score (enter NT for not tested) Directions: Say to the patient, "I am going to read you a series of 10 letters. Whenever you hear the letter 'A,' indicate by squeezing my hand." Read letters from the following letter list in a normal tone. **S A V E A H A A R T** Scoring: Errors are counted when patient fails to squeeze on the letter "A" and when the patient squeezes on any letter other than "A."	**Score (out of 10): _____**	
2B: ASE Pictures: record score (enter NT for not tested) Directions are included on the picture packets.	**Score (out of 10): _____**	
Feature 3: Disorganized Thinking Positive if the combined score is less than 4	**Positive**	**Negative**
3A: Yes/No Questions (Use either Set A or Set B, alternate on consecutive days if necessary): **Set A** 1. Will a stone float on water? 2. Are there fish in the sea? 3. Does one pound weigh more than two pounds? 4. Can you use a hammer to pound a nail? **Set B** 1. Will a leaf float on water? 2. Are there elephants in the sea? 3. Do two pounds weigh more than one pound? 4. Can you use a hammer to cut wood? **Score** ___ (Patient earns 1 point for each correct answer out of 4) **3B: Command** Say to patient: "Hold up this many fingers" (Examiner holds two fingers in front of patient) "Now do the same thing with the other hand" (Not repeating the number of fingers). *If pt is unable to move both arms, for the second part of the command ask patient "Add one more finger." **Score**___ (Patient earns 1 point if able to successfully complete the entire command)	**Combined Score (3A+3B): _____ (out of 5)**	
Feature 4: Altered Level of Consciousness Positive if the Actual RASS score is anything other than "0" (zero)	**Positive**	**Negative**
Overall CAM-ICU (Features 1 and 2 and either Feature 3 or 4):	**Positive**	**Negative**

FIGURE 33–3. Worksheet from the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) training manual

ASE=Attention Screening Examination; GCS=Glasgow Coma Score; pt=patient; RASS=Richmond Agitation-Sedation Scale.

Source. Ely EW: *The Confusion Assessment Method for the Intensive Care Unit (CAM-ICU): Training Manual.* Copyright 2002, E. Wesley Ely, M.D., M.P.H. and Vanderbilt University, all rights reserved. Used with permission.

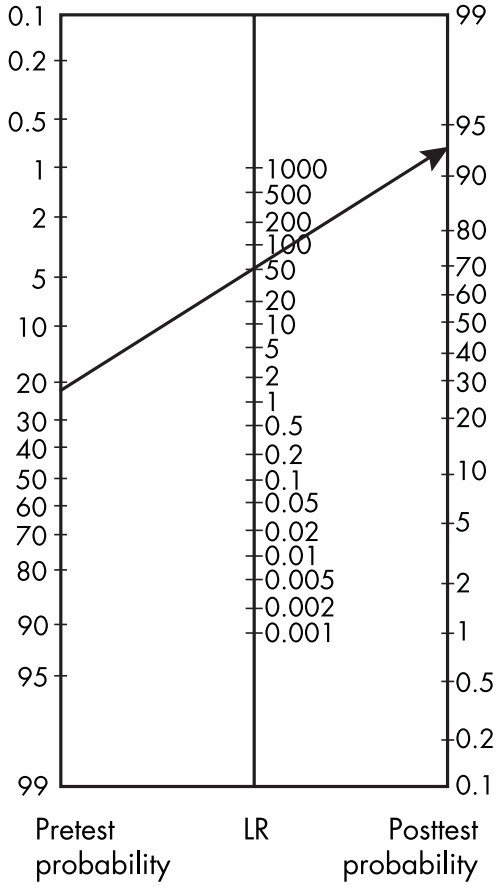


FIGURE 33–4. Likelihood ratio (LR) nomogram for posttest probability of delirium.

The possibility of medication-induced mental status changes as a result of the opiates and intermittent benzodiazepines that he was receiving appeared likely. The possibility of alcohol use and withdrawal were especially of concern because his alcohol and substance use had not been well documented in the record, and he was noted to have been trembling and diaphoretic by his family. He had several possible reasons for tachycardia, but his pulse remained elevated despite correction of these factors and he continued to exhibit tremor and diaphoresis.

Worsening of his cardiac and respiratory function was a possible contributor, given his previous smoking history and known cardiac disease, the tachyarrythmias noted, and his pulmonary congestion. His atrial fibrillation was now more optimally rate controlled, his breathing and chest X ray were improved, and his ECG and cardiac enzymes were unremarkable. He had an echocardiogram performed, which had shown normal left ventricular (LV) size with LV systolic function at lower limits of normal and inferior wall hypokinesis, and mildly thickened mitral valve with mild mitral regurgitation. The right ventricle was of normal size, with normal ejection fraction. Both atria were severely dilated. Findings were unchanged compared with the prior study from 6 months prior to admission. He was no longer tachypneic, his oxygen saturation and arterial blood gas were unremarkable. He had no unilateral lower extremity edema or pain with palpation of his deep leg veins, he had no hemoptysis, and his acute hypoxia seemed to be better accounted for by congestive heart failure than by a pulmonary embolus. However, he was recently immobilized and had surgery for his cancer, and he was tachycardic, which did indicate a moderate pretest probability by the simplified Wells criteria for a pulmonary embolus (Chunilal et al. 2003).

The possibility of infection was present given his recent fever. Several possible sites could be considered. He had recently had his procedure on his larynx, he also had complications of his gastrostomy tube

extravasation into his peritoneal space, and he had respiratory difficulties and abnormal urinalysis. His fever, together with his having been ventilated, with new infiltrates on chest X ray, could suggest a pulmonary infection. The absence of a white count or shift and the absence of purulent sputum could not exclude the possibility of infection (Klompas 2007). He did have other reasons to have a pulmonary infiltrate and did seem to be responding well to treatment. The bladder catheter and central venous catheters were other possible sources of infection. Blood, sputum, and urine cultures had been sent to the lab, and the patient had been started on antibiotics.

A cerebrovascular accident was possible given his recent unanticoagulated atrial fibrillation; however, his neurologic exam was unremarkable, with the exception of his altered mental state, and a head computed tomographic image obtained prior to our consultation by the medical service was unremarkable. We felt this was less likely.

The patient's family denied he had any significant visual or hearing impairment or any history of depression, dementia, or other psychiatric illnesses.

The most important possibility that did not seem to have been adequately considered was alcohol withdrawal. We thus asked the family if the patient consumed alcohol, and they indicated that he did not like to talk about this because he knew the doctors did not approve of this given his throat cancer. They did acknowledge, when informed that this could have important and possibly life-threatening consequences, that he had consumed alcohol in addition to cigarettes for many years and that he continued up until the night prior to hospitalization to consume at least three to four cocktails every evening, and he was known at times to have more. This had not previously been noted in the history. The patient remained tachycardic, which had reasonably been attributed to the many factors cited previously plus postoperative pain and discomfort. His recent tremulousness and disturbance in mental state had been attributed to anxiety and emotional upset. Alcohol withdrawal delirium rose to be included as a possible contributing factor that would be likely to require a change in the management of the patient. We considered it difficult to determine, however, whether all of the symptoms of autonomic hyperarousal could be because of alcohol withdrawal but felt that the risk was high enough that it should be considered. The other factors were being carefully considered in the management of the patient, and treatment decisions were being made accordingly.

We discussed our impression with the team and the patient's family. The patient was assessed as lacking the capacity to understand his current condition, the treatment recommendations, and alternatives, including the consequences of not receiving treatment, and thus we felt he could not provide informed consent for his care. We thus recommended that the patient's wife act as a surrogate for her husband. We discussed the general guidelines suggested for the management of delirium (British Geriatrics Society 2006), noting that the patient was hyperactive at times and at risk for falls. Shortly after we were consulted he had attempted to pull out his intravenous lines. We felt that this hyperactive delirium and the risk for the complications of alcohol withdrawal were going to require pharmacological intervention. We advised this in addition to the environmental interventions suggested by the guideline. This guideline did not cover alcohol withdrawal, and thus treatment recommendations for the pharmacologic management of alcohol withdrawal were made in reference to a previously appraised systematic review of reasonable quality (Mayo-Smith et al. 2004) on alcohol withdrawal delirium. It was noted that injudicious use of symptom-triggered withdrawal protocols could be associated with harms in patients unable to communicate well who had alternative causes of their symptoms that might be consistent with those symptoms produced by alcohol withdrawal (Hecksel et al. 2008). Making treatment recommendations for this type of case is more challenging than for the typical alcohol withdrawal delirium case. The Clinical Institute Withdrawal Assessment for Alcohol was not considered an appropriate instrument for following the patient's progress given his inability to communicate well and was not recommended. Treatment of delirium with benzodiazepines has been considered to be a risk factor for worsening delirium by one small randomized controlled trial of modest quality in hospitalized AIDS patients (Breitbart et al. 1996) and is often cited as a possible causal risk factor based upon observational studies of varying quality in general medical patients (British Geriatrics Society 2006).

We opted to recommend that the team follow the patient's vital signs and CAM-ICU ratings each shift and to use intravenous lorazepam with the goal of attaining light sedation to the point of a Richmond

Agitation-Sedation Scale score of –1 during the day, with sleep the goal overnight, consistent with the recommendations in the systematic review by Mayo-Smith (Mayo-Smith et al. 2004). We felt that despite the CAM's dichotomous nature, the RASS score would be valuable along with the vital signs and other signs of autonomic hyperactivity in continuing to guide our approach, which was focused on reducing the risks of hyperactive delirium (line pulling, getting out of bed, taking things orally despite NPO status) and the effects of alcohol withdrawal. Because other possible causes of delirium were present, we discussed additional use of a neuroleptic agent (Lonergan et al. 2007) if the intervention with the benzodiazepine was ineffective and the patient remained severely agitated and at risk for pulling lines or falling from his bed. We opted for a watchful waiting approach given the patient's prolonged QTc of 470 msec (Ray et al. 2009) and opted to see if he might respond to the benzodiazepine protocol first. The patient's wife provided informed consent, and treatment along these lines was initiated.

The patient's course gradually improved. The remainder of his medical course was relatively unremarkable. He suffered no obvious motor seizures or worsening of his medical condition. He was intermittently hypoactively delirious by CAM-ICU criteria for several more days. He was no longer hyperactively delirious, however, and was significantly calmer, with markedly decreased agitation, tremor, diaphoresis, and tachycardia. He required no neuroleptics. His mental status improved significantly to being CAM-ICU negative consistently after 1 week, although residual cognitive symptoms remained at 2 weeks, at which time the patient had been completely titrated off of all benzodiazepines and had recovered medically sufficiently to be discharged from the hospital. We recommended that he be discharged home with additional in-home nursing support made available to him. We noted, on the basis of the limited literature in this area, that his prognosis could possibly include longer-term cognitive difficulties that might be significant and that might require continued supportive resources for him and his wife (Gordon et al. 2004).

• Now take a moment to review the list of three to five diagnoses that you wrote down at the start of the case presentation. *Hindsight bias* occurs when we overestimate the probability of a given diagnosis when that diagnosis has been established previously (Bornstein and Emler 2001). Things often appear quite obvious in looking backwards at a case. We hope that the previous case presentation was helpful in discussing some of the common biases affecting our judgment and decision making and that you find some things in the discussion of use in your practice.

Summary

- Selecting a treatment intervention strongly supported by high-quality randomized controlled trials is unlikely to benefit the patient and may cause harm if the diagnosis is incorrect. Treating this patient for major depressive disorder or even for delirium with a neuroleptic agent given his cardiac history, prolonged QTc, and likely alcohol withdrawal could have resulted in a poorer outcome.
- The EBM approach to diagnosis is still in its infancy. The literature is sparse, and there are not yet large repositories of systematic reviews of diagnostic accuracy or precision available to the clinician.
- Various approaches to generating a differential diagnosis can be considered. One might consider combining approaches to avoid common biases that result in diagnostic error: a broad differential diagnosis can be generated (possibilistic approach) that weights more heavily those causes that cannot be missed (prognostic approach) as well as those that are highly likely in that practice setting (probabilistic approach) that can be treated resulting in better patient outcomes (pragmatic approach).
- The *DSM-IV-TR Handbook of Differential Diagnosis* (First et al. 2002) can be used to quickly generate a differential diagnosis likely to be consistent with those considered in randomized controlled trials in psychiatry.
- Carefully considering estimates of the pretest probability of conditions that are pertinent in the differential diagnosis in the setting in which the patient is encountered is important in making a valid diagnosis and avoiding representativeness bias.
- Preappraisal of bedside diagnostic procedures important in clinical settings for the common diagnoses encountered may increase the speed and confidence in the diagnoses assigned in those settings.
- The STARD checklist or other structured appraisal tools for diagnostic quality can be used to

assist practitioners in identifying and assembling a toolkit of valid and reliable appraisal instruments for approaching the common differential diagnoses occurring in their practice settings.
- The operating characteristics of various diagnostic tests can be considered together with the pretest probabilities of the conditions being considered to estimate the posttest probability of that condition being present, because tests are not perfect and representativeness bias can strike again.
- Good clinical judgment is necessary at every step.

References

AGREE Collaboration: Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 12:18–23, 2003

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000

Ardern M, Mayou R, Feldman E, et al: Cognitive impairment in the elderly medically ill: how often is it missed? Int J Geriatr Psychiatry 8:929–937, 1993

Bandolier: Evidence and diagnostics. February 2002. Available at: http://www.medicine.ox.ac.uk/bandolier/booth/diagnos/Diagessy.html. Accessed May 18, 2009

Boland RJ, Diaz S, Lamdan RM, et al: Overdiagnosis of depression in the general hospital. Gen Hosp Psychiatry 18:28–35, 1996

Bornstein BH, Emler AC: Rationality in medical decision making: a review of the literature on doctors' decision-making biases. J Eval Clin Pract 7:97–107, 2001

Bossuyt PM, Reitsma JB, Bruns DE, et al: The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Ann Intern Med 138:W1–W12, 2003a

Bossuyt PM, Reitsma JB, Bruns DE, et al: Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. Ann Intern Med 138:40–44, 2003b

Boustani M, Peterson B, Hanson L, et al: Screening for dementia in primary care: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 138:927–937, 2003

Breitbart W, Marotta R, Platt MM, et al: A double-blind trial of haloperidol, chlorpromazine, and lorazepam in the treatment of delirium in hospitalized AIDS patients. Am J Psychiatry 153:231–237, 1996

British Geriatrics Society: Guidelines for the prevention, diagnosis and management of delirium in older people: concise guidance to good practice series. January 2006. Available at: http://www.bgs.org.uk/Publications/Clinical%20Guidelines/clinical_1-2_fulldelirium.htm. Accessed May 28, 2009

Casscells W, Schoenberger A, Graboys TB: Interpretation by physicians of clinical laboratory results. N Engl J Med 299:999–1001, 1978

Chodosh J, Petitti DB, Elliott M, et al: Physician recognition of cognitive impairment: evaluating the need for improvement. J Am Geriatr Soc 52:1051–1059, 2004

Chunilal SD, Eikelboom JW, Attia J, et al: Does this patient have pulmonary embolism? JAMA 290:2849–2858, 2003

Deeks JJ, Altman DG: Diagnostic tests 4: likelihood ratios. BMJ 329:168–169, 2004

Devlin JW, Fong JJ, Fraser GL, et al: Delirium assessment in the critically ill. Intensive Care Med 33:929–940, 2007

Ely EW, Inouye SK, Bernard GR, et al: Delirium in mechanically ventilated patients: validity and reliability of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). JAMA 286:2703–2710, 2001a

Ely EW, Margolin R, Francis J, et al: Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). Crit Care Med 29:1370–1379, 2001b

Ely EW, Shintani A, Truman B, et al: Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. JAMA 291:1753–1762, 2004

Fagan TJ: Nomogram for Bayes theorem (letter). N Engl J Med 293:257, 1975

Farrell KR, Ganzini L: Misdiagnosing delirium as depression in medically ill elderly patients. Arch Intern Med 155:2459–2464, 1995

First MB, Frances A, Pincus HA: DSM-IV-TR Handbook of Differential Diagnosis. Washington, DC, American Psychiatric Publishing, 2002

Ghosh AK, Ghosh K, Erwin PJ: Do medical students and physicians understand probability? QJM 97:53–55, 2004

Glasziou P: Which methods for bedside Bayes? Evid Based Med 6:164–166, 2001

Gordon SM, Jackson JC, Ely EW, et al: Clinical identification of cognitive impairment in ICU survivors: insights for intensivists. Intensive Care Med 30:1997–2008, 2004

Gustafson Y, Brännström B, Norberg A, et al: Underdiagnosis and poor documentation of acute confusional states in elderly hip fracture patients. J Am Geriatr Soc 39:760–765, 1991

Harwood DM, Hope T, Jacoby R: Cognitive impairment in medical inpatients. I: Screening for dementia—is history better than mental state? Age Ageing 26:31–35, 1997

Hecksel KA, Bostwick JM, Jaeger TM, et al: Inappropriate use of symptom-triggered therapy for alcohol withdrawal in the general hospital. Mayo Clin Proc 83:274–279, 2008

Hodkinson HM: Evaluation of a mental test score for assessment of mental impairment in the elderly. Age Ageing 1:233–238, 1972

Holsinger T, Deveau J, Boustani M, et al: Does this patient have dementia? JAMA 297:2391–2404, 2007

Inouye SK: Delirium in hospitalized older patients: recognition and risk factors. J Geriatr Psychiatry Neurol 11:118–125; discussion 157–158, 1998

Inouye SK, van Dyck CH, Alessi CA, et al: Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med 113:941–948, 1990

Inouye SK, Foreman MD, Mion LC, et al: Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings. Arch Intern Med 161:2467–2473, 2001

Jorm AF: The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): a review. Int Psychogeriatr 16:275–293, 2004

Kahneman D, Slovic P, Tversky A (eds): Judgment Under Uncertainty: Heuristics and Biases. Cambridge, UK, Cambridge University Press, 1982

Klompas M: Does this patient have ventilator-associated pneumonia? JAMA 297:1583–1593, 2007

Kraemer HC, Kazdin AE, Offord DR, et al: Coming to terms with the terms of risk. Arch Gen Psychiatry 54:337–343, 1997

Lonergan E, Britton AM, Luxenberg J, et al: Antipsychotics for delirium. Cochrane Database Syst Rev (2):CD005594, 2007

Lyman GH, Balducci L: Overestimation of test effects in clinical judgment. J Cancer Educ 8:297–307, 1993

Lyman GH, Balducci L: The effect of changing disease risk on clinical reasoning. J Gen Intern Med 9:488–495, 1994

Mascola AJ: Guided mentorship in evidence-based medicine for psychiatry: a pilot cohort study supporting a promising method of real-time clinical instruction. Acad Psychiatry 32:475–483, 2008

Mayo-Smith MF, Beecher LH, Fischer TL, et al: Management of alcohol withdrawal delirium: an evidence-based practice guideline. Arch Intern Med 164:1405–1412, 2004

McCusker J, Cole M, Abrahamowicz M, et al: Delirium predicts 12-month mortality. Arch Intern Med 162:457–463, 2002

McGee S: Evidence-Based Physical Diagnosis. Philadelphia, PA, WB Saunders, 2007

Nicholas LM, Lindsey BA: Delirium presenting with symptoms of depression. Psychosomatics 36:471–479, 1995

Olafsdottir M, Skoog I, Marcusson J: Detection of dementia in primary care: the Linkoping study. Dement Geriatr Cogn Disord 11: 223–229, 2000

Pauker S, Kassirer J: The threshold approach to clinical decision making. N Engl J Med 302:1109–1117, 1980

Pisani MA, Inouye SK, McNicoll L, et al: Screening for preexisting cognitive impairment in older intensive care unit patients: use of proxy assessment. J Am Geriatr Soc 51:689–693, 2003a

Pisani MA, Redlich C, McNicoll L, et al: Underrecognition of preexisting cognitive impairment by physicians in older ICU patients. Chest 124:2267–2274, 2003b

Ray WA, Chung CP, Murray KT, et al: Atypical Antipsychotic Drugs and the Risk of Sudden Cardiac Death. N Engl J Med 360:225–235, 2009

Richardson WS: We should overcome the barriers to evidence-based clinical diagnosis! J Clin Epidemiol 60:217–227, 2007

Richardson WS, Wilson MC, Nishikawa J, et al: The well-built clinical question: a key to evidence-based decisions. ACP J Club 123:A12–A3, 1995

Richardson WS, Wilson MC, Guyatt GH, et al: Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. Evidence-Based Medicine Working Group. JAMA 281:1214–1219, 1999

Richardson WS, Glasziou P, Polashenski WA, et al: A new arrival: evidence about differential diagnosis. ACP J Club, 133:A11–A12, 2000a

Richardson WS, Wilson MC, Williams JW Jr, et al: Users' guides to the medical literature: XXIV. How to use an article on the clinical manifestations of disease. Evidence-Based Medicine Working Group. JAMA 284:869–875, 2000b

Richardson WS, Polashenski WA, Robbins BW: Could our pretest probabilities become evidence based? A prospective survey of hospital practice. J Gen Intern Med 18:203–208, 2003

Roca RP, Klein LE, Kirby SM, et al: Recognition of dementia among medical patients. Arch Intern Med 144:73–75, 1984

Rothman AJ, Kiviniemi MT: Treating people with information: an analysis and review of approaches to communicating health risk information. J Natl Cancer Inst Monogr 25:44–51, 1999

Sackett DL: The rational clinical examination. A primer on the precision and accuracy of the clinical examination. JAMA 267:2638–2644, 1992

Sackett DL, Straus SE: Finding and applying evidence during clinical rounds: the "evidence cart." JAMA 280:1336–1338, 1998

Siddiqi N, House AO, Homes JD: Occurrence and outcome of delirium in medical in-patients: a systematic literature review. Age Ageing 35:350–364, 2006

Steurer J, Fischer JE, Bachmann LM, et al: Communicating accuracy of tests to general practitioners: a controlled study. BMJ 324: 824–826, 2002

Streiner DL: Diagnosing tests: using and misusing diagnostic and screening tests. J Pers Assess 81:209–219, 2003

Swigart SE, Kishi Y, Thurber S, et al: Misdiagnosed delirium in patient referrals to a university-based hospital psychiatry department. Psychosomatics 49:104–108, 2008

Valcour VG, Masaki KH, Curb JD, et al: The detection of dementia in the primary care setting. Arch Intern Med 160:2964–2968, 2000

Whiting PF, Weswood ME, Rutjes AW, et al: Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol 6:9, 2006

34

Alcohol Dependence Treated by a Psychiatry Resident

Violeta Tan, M.D.

Setting

This patient was treated while I was on a 6-month rotation during my fourth year as an adult psychiatry resident in a university outpatient clinic. The clinic serves a mostly insured population with a broad ethnic diversity. The focus is on short-term therapies using a combination of psychotherapy and psychopharmacology.

Illustration

- Use of the American Psychiatric Association (APA) *Practice Guideline for the Treatment of Patients With Substance Use Disorders* (2006)
- Use of relapse prevention strategies
- Use of motivational interviewing principles
- Use of psychotherapy termination recommendations

Chief Complaint

Ms. M.T. is a 35-year-old divorced female engineer who presents with a chief complaint of depression.

Present Illness

Ms. M.T. had been seen by two previous adult psychiatry residents over a span of 8 months prior to my first meeting her. She initially presented to the clinic with worsening depressive symptoms following her discontinuation of citalopram. Psychosocial stressors included financial negotiations related to her divorce and her son's living with her ex-husband. She was restarted on citalopram, which was eventually augmented with bupropion.

She had been drinking approximately a quarter to a half of a bottle of wine per night over the past 3 years. She noted diminished effect with continued use of the same amount and admitted the amount had increased over a longer period than she had intended. Furthermore, she found herself often turning to alcohol instead of engaging in her usual recreational or social activities, losing touch with friends in the process. She reported feeling the need to cut down on her drinking, and, although her previous mental health care providers had encouraged her to stop drinking and to attend Alcoholics Anonymous (AA), she resisted their recommendations for months. However, the day after termination with her previous psychiatrist, before beginning therapy with me, the patient did attend AA and decided to abstain.

On our first meeting, Ms. M.T. was proud of herself for having attended AA. Yet she also admitted to feelings of grief and loss in not having alcohol available any longer. She shared that she often used alcohol as a coping mechanism to quell feelings of sadness and isolation. Further review of psychiatric symptoms revealed anhedonia, sleep disturbances with awakenings in the middle of the night, energy loss, and concentration difficulties. Her appetite was better than usual and she denied thoughts of self-harm. In addition to her current symptoms, she provided a clear past history of depression in the

absence of substances. Ms. M.T. reported continuing depression over her failed marriage, financial burdens, and worries about future interactions with her ex-husband. She furthermore described a longing to return to her hometown on the East Coast where her family resided yet also felt obligated to remain in her current location for her 13-year-old son, who was in his last year of grade school.

Other Significant Findings From Assessment

Ms. M.T. denied tobacco use. She drank about two cups of a caffeinated beverage each day. She did not exercise regularly and was unsatisfied with her weight. Family history was significant for alcoholism and anxiety among her grandparents, parents, and siblings. Siblings and grandparents also suffered from depression.

DSM-IV-TR Diagnosis

Axis I	Major depressive disorder, recurrent Alcohol dependence
Axis II	None
Axis III	Asthma
Axis IV	Limited social support, strain with ex-husband, living circumstances, financial burden
Axis V	GAF score: 60

Treatment Plan Considerations

Problems

The patient's main problems were depression and alcohol dependence. In her medical history form, she had written that her main reasons for seeking treatment were "depression, wanting to make changes, and needing support." She also hoped to maintain abstinence but seemed more concerned with her mood and felt her drinking was under control.

Secondary issues were related to interpersonal relationships, which partly fueled her depressed mood. She felt isolated and longed to return to her family on the East Coast, yet was tied to her current residence for her son's benefit. Beyond her current interpersonal issues and future concerns, her past divorce haunted her.

Selecting a Guideline

My supervisor referred me to the APA *Practice Guideline for the Treatment of Patients With Substance Use Disorders* (2006). The general treatment principles are broken into eight steps followed by psychiatric management, each with detailed recommendations. By the time I first met the patient, I felt that several of the APA's recommended steps were already achieved by my predecessors. The goals of treatment, as suggested in the guideline, including motivating the patient to change, reducing substance use or achieving complete abstinence, and improving social functioning, were already under way. Because the patient was abstinent when I first met her and had started attending AA, I anticipated I would focus on the relapse prevention section of the guideline (Table 34–1). In reading about counseling for alcohol-related disorders, I knew much of the emphasis would be on how to build a lifestyle free of alcohol (Sadock and Sadock 2003). I expected to use relapse prevention strategies and motivational interviewing while also encouraging support groups such as AA.

To address her depression, I decided to monitor her mood and substance use weekly via general assessment charts provided by the clinic (Figures 34–1 through 34–3). Note that I began to use the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}) (University of Pittsburgh Epidemiology Data Center 2009) for assessment later in the course of therapy.

With these considerations, I discussed an initial treatment plan and approach with the patient, outlined in Table 34–2.

I continued the patient's regimen of citalopram 60 mg every day (qd) and bupropion XL 300 mg qd. Ms. M.T. denied any side effects. I scheduled her follow-up visit in 1 week.

Course

Visit 2 (Week 2)

Ms. M.T. continued to abstain from alcohol, now 9 days sober, although she admitted to having some difficulty with cravings. She had not attended AA in the past week. I explained that I would begin monitoring her depression, drinking, and craving trends from this session onward. This day, she rated her depression at 7/10 (Figure 34–1). Her interpersonal

TABLE 34–1. Treatment strategy for alcohol dependence

Psychiatric treatment	Followed
Use relapse prevention strategies:	√
Help the patient anticipate and avoid drug-related cues (e.g., instruct the patient to avoid drug-using peers).	√
Decrease access to abusable substances.	√
Train the patient to self-monitor states associated with increased craving.	√
Use contingency contracting (e.g., set up positive and negative reinforcements in advance).	√
Teach desensitization and relaxation techniques to reduce the power of substance-related stimuli.	√
Help patients develop alternative, nonchemical coping responses.	√
Provide social skills training.	√
Provide positive feedback for the patient's successes, even if relapse does occur.	√
Analyze relapses and periods of sobriety from functional and behavioral standpoints and modify the treatment plan, including psychotherapy, accordingly.	√

Source. Adapted from American Psychiatric Association: *Practice Guideline for the Treatment of Patients With Substance Use Disorders*. 2006. Available at: http://www.psychiatryonline.com/pracGuide/loadGuidelinePdf.aspx?file=SUD2ePG_04-28-06. Accessed June 1, 2009.

stressors were related to forming a relationship with a man in the East and having concerns over sustaining it given the distance. Even more, Ms. M.T. felt her son disapproved of this developing bond. She was struggling with a desire to fulfill her own needs versus the needs of her son—wanting to move East where her family and new boyfriend were living yet feeling a responsibility to remain in California at least until her son completed the eighth grade. She revealed unmet fantasies of having more children to create the picture of a family she had envisioned with a stable marriage. During this session, I primarily used supportive and interpersonal therapy techniques in addressing her relationship issues. I recognized Ms. M.T.'s 9 days of sobriety and also encouraged her to attend AA.

TABLE 34–2. Initial treatment goals, measures, and methods of case

Treatment goal	Measure	Method
Mood improvement and stabilization	Weekly quantitative assessment of mood STAR*D QIDS-SR_{16}[a]	Pharmacotherapy Techniques from interpersonal therapy and supportive therapy
Reduce depressive symptoms by 50% by week 8	Weekly quantitative assessment of mood STAR*D QIDS-SR_{16}[a]	Pharmacotherapy Techniques from interpersonal therapy and supportive therapy
Abstain from alcohol	Self-report (episodes/week)	Relapse prevention Motivational interviewing Encourage attending Alcoholics Anonymous
Improved interpersonal relationships	Self-report (qualitative)	Techniques from interpersonal therapy and supportive therapy

Note. QIDS=Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}); STAR*D=Sequenced Treatment Alternatives to Relieve Depression.

[a]Added at visit 10 (week 12).

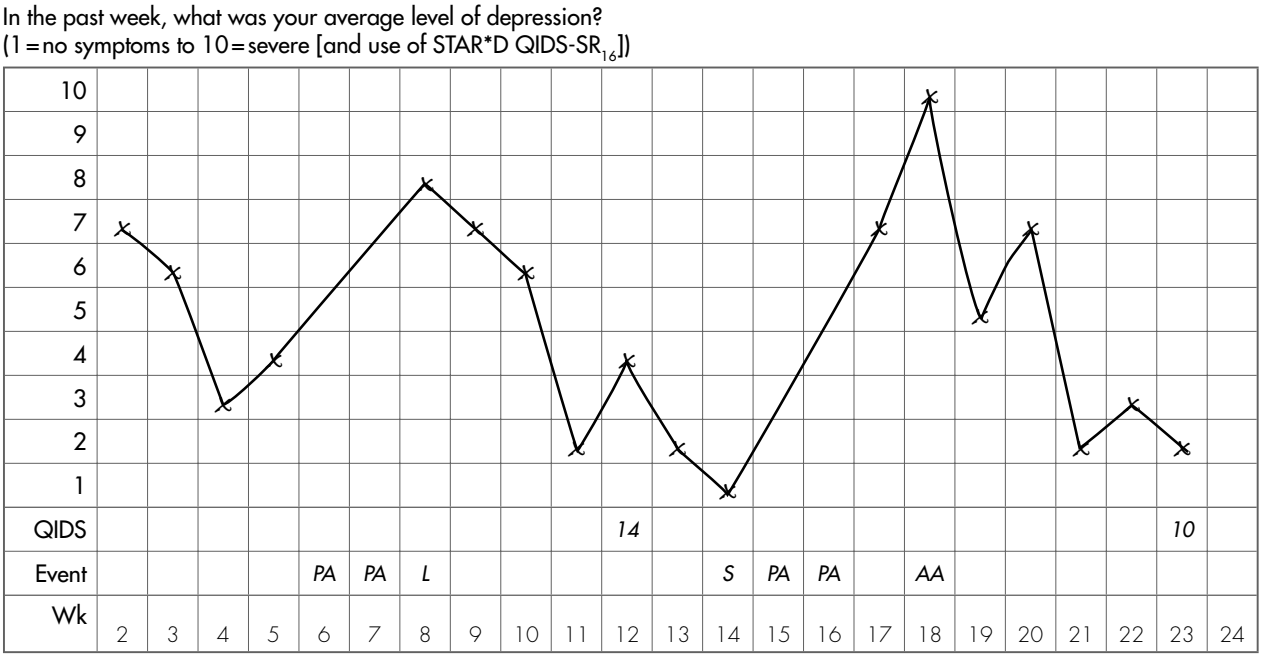


FIGURE 34–1. Depression trends.

AA=author away and no appointment; L=lapse; PA=patient away and no appointment; S=session addressing alcohol as a problem. QIDS=Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}); STAR*D=Sequenced Treatment Alternatives to Relieve Depression.

Visit 3 (Week 3)

Ms. M.T.'s mood remained relatively unchanged from her last visit, with depression rated at 6/10. She successfully abstained for yet another week, but described significant, strong cravings triggered by walking down a wine-filled aisle at the grocery store. Related to the APA guideline suggestion of anticipating and avoiding drug-related cues as a relapse prevention strategy, I recommended that she self-monitor states associated with increased craving. I spent much of this session gathering the patient's past alcohol history to better understand her present picture. This was also in line with the second step of the general treatment principles of the APA guideline in assessing a patient (Table 34–3). Although assessment of the patient's history of use had been done in previous sessions by her prior therapists, I thought it important and worthwhile to conduct my own assessment—both to offer me a better understanding of the patient's current situation and for the patient to reassess her goals.

The patient described growing up in a household where alcohol was the norm, her parents drinking daily. She began drinking at the age of 13, often mixing drinks found in her and her friends' homes. Her drinking progressed into college and beyond, with

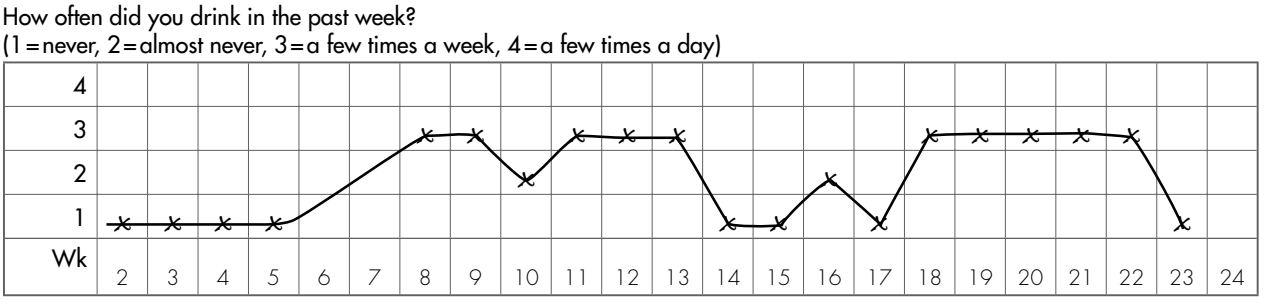


FIGURE 34–2. Alcohol use.

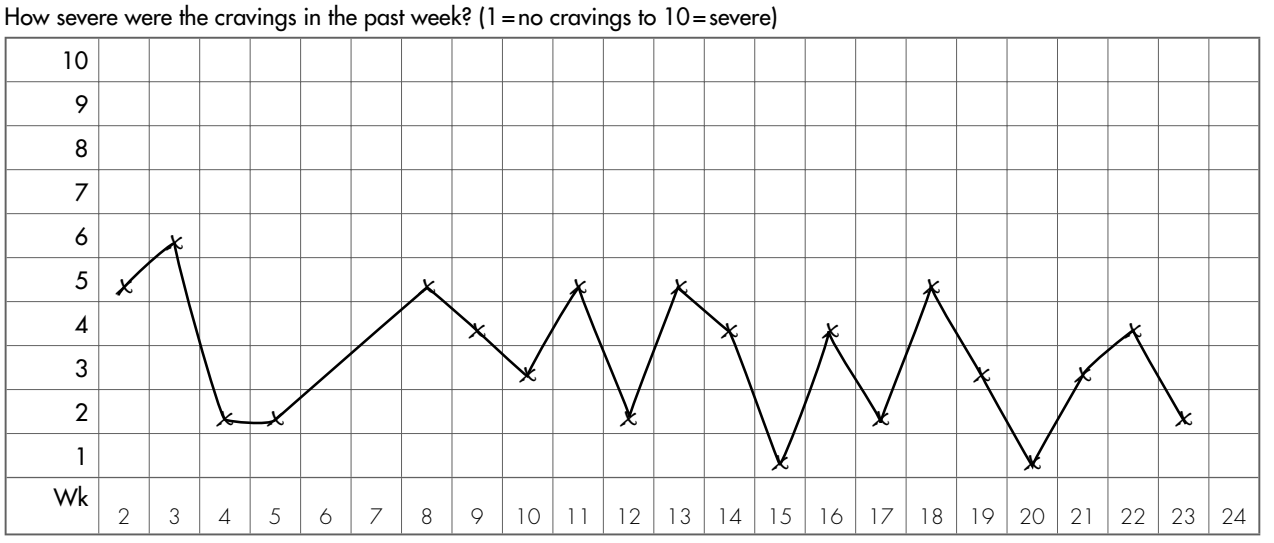


FIGURE 34–3. Cravings.

brief periods of abstinence. Her longest period of sobriety was about 7 years while she was with her ex-husband, who did not drink. After their separation, she began drinking off and on, oftentimes returning to alcohol under stress.

I focused on events that she could recall that led to her drinking, identifying her triggers in the process. In the past year, her drinking was motivated by a desire to numb out personal problems related to her ex-husband. By the end of this session, we had gathered a handful of triggers to be mindful of for the future, most notably including the sight of alcoholic beverages and periods of depression and isolation.

Ms. M.T. was preparing for a trip to the East Coast to visit family and expressed concerns of relapse in this familiar environment with many potential alcohol-related triggers. In following the relapse prevention strategies (Table 34–1), we discussed alternative, nonchemical coping responses, which included attending AA on the East Coast, informing family in advance about her goals, using nonalcoholic beer as a substitute, and finding alternate activities such as art. During this session, I utilized several of the APA's relapse prevention strategies (Table 34–1), with some points addressed or expanded on in future encounters.

Visits 4–5 (Weeks 4–5)

Ms. M.T.'s mood was notably improved the next couple visits, rated at 3–4/10. This was partly attributed to improved relations with her boyfriend and son. She had visited her boyfriend recently, which eased some of her loneliness. She also spoke with her son about her dilemma of wanting to return to the east yet not wanting to remove him from his comfortable California network of friends. Although there was no firm resolution with this issue, she was relieved after sharing her struggles with the individual she felt it would most affect.

She continued to successfully abstain from alcohol, finding support in her new boyfriend, who I discovered had a history of alcohol dependence and who was, at that time, abstaining. Ms. M.T.'s cravings were also diminished, rated at 2/10. We continued to discuss strategies for preventing relapse, particularly for her upcoming visit back home. Another trigger for cravings was identified—that of boredom leading to depressed mood. To prevent relapse, she engaged herself in various activities after work, including painting, piano, and swimming.

We reinforced the various strategies and coping skills for her to use during her 2-week trip to the East Coast. She was advised to seek AA while there and to return to the clinic when she arrived back in California.

Visit 6 (Week 8)

Ms. M.T. returned from her trip, where she had successfully remained sober. However, upon her return to California, she lapsed, experiencing loneliness and boredom, which led to her drinking leftover al-

TABLE 34–3. Assessment of alcohol history strategy

Assessment	Followed
Obtain information from the patient and, with the patient's permission, from collateral sources (e.g., available family members, friends, current and past treaters, employers) as appropriate	√
Obtain detailed history of the patient's past and present use of substances, including:	
Types of substance used (including nicotine, caffeine, prescribed and over-the-counter medications) and whether multiple substances are used in combination	√
Mode of onset, quantity, frequency, duration, route of administration, and pattern and circumstances of substance use (e.g., where, with whom)	√
Timing and amount of most recent use	√
Degree of associated intoxication, withdrawal, and subjective effects of all substances used	√
History of prior substance use treatments (e.g., settings, context, modalities, duration, and adherence), efforts to stop substance use, and outcomes (e.g., duration of abstinence, subsequent substance use, reasons for relapse, social and occupational functioning achieved)	√
Current readiness to change, including:	
Awareness of substance use as a problem	√
Plans for ceasing substance use	√
Motivations for substance use, including desired effects	√
Barriers to treatment and to abstinence	√
Expectations and preferences for future treatment	√
Effects on cognitive, psychological, behavioral, social, occupational, and physiological functioning	√
General medical and psychiatric history and physical and mental status examination	√
Family history of substance use or psychiatric disorder	√
Social history (including family and peer relationships, financial problems, and legal problems) and psychosocial supports (including influence of close friends or other household members to support or undermine past efforts at abstinence)	√
Educational and occupational history, including school or vocational adjustment and identification of occupations at increased risk of substance use	√

Source. Adapted from American Psychiatric Association: *Practice Guideline for the Treatment of Patients With Substance Use Disorders*, 2nd Edition. 2006. Available at: http://www.psychiatryonline.com/pracGuide/loadGuidelinePdf.aspx?file=SUD2ePG_04-28-06. Accessed June 1, 2009.

cohol in her home. Her depression was rated higher than when we first met. We discussed reasons for her lapse, and she identified missing her family and boyfriend. We agreed that we had planned adequately for her trip to the east but had not anticipated the difficulty she might have returning to California.

In addressing where to go at this point, Ms. M.T. was ambivalent. She expressed a desire to limit her drinking to weekends, giving up her goal of complete abstinence but also hoping to return to that goal eventually. Although I was already using motivational interviewing, I knew this clinical method would be especially relevant in exploring and resolving the patient's ambivalence. We set a goal for her to contact her former neighbor, who also was recovering from alcohol dependence, to attend AA with her.

After this lapse, I found myself searching for more in-depth relapse prevention guidelines. I logged onto the National Guideline Clearinghouse at www.guideline.gov, but the recommendations I did find were strategies I had already been using or more related to screening for substance use. Therefore, I turned to researching articles online and seeking advice from supervisors. I presented the case during a conference to a group of psychiatrists and psychologists, seeking their opinions. I also video-

taped my sessions and showed them to my supervisor for input.

I came across an article online with suggestions on how to prevent a lapse from becoming a relapse, the main message being to view the lapse not as a failure but as a reminder flag to be vigilant about one's behavior—to learn from the lapse rather than using it as an excuse to relapse completely (Hunt 2007). From my last session with Ms. M.T., I feared she might already be viewing her lapse as an excuse to continue drinking—she was considering a new goal that would allow her to drink on weekends instead of completely abstaining. Hence, prior to her next visit, I called Ms. M.T., validating her efforts and achievements of the past several weeks and giving her affirmation that her lapse did not need to become a relapse. The online suggestion allowed me to present to Ms. M.T. another way of viewing her setback that aided in maintaining hope and self-esteem. Ms. M.T. expressed her appreciation for the call.

Visits 7–11 (Weeks 9–13)

Over the next several sessions, my patient continued to drink approximately two to three glasses of wine two to three times per week. Beyond using alcohol as a coping mechanism for depression and isolation, I recognized that she began to use it as a reward for accomplishing things such as exercising or even going to AA. During week 9, I discovered that she did not feel alcohol was that problematic for her, believing that she could control her use. I realized the initial steps in the APA guideline, which we had once gone through, needed to be reassessed, particularly with current readiness to change (Table 34–3). Ms. M.T. felt that because she had lapsed, she could indulge in alcohol before abstaining again. Further complicating the picture, her boyfriend, who had also lapsed, stayed with her for a period and encouraged her to drink to celebrate their reunion.

Throughout these sessions, we discussed alternate forms of reward, incorporating other activities into her life such as exercise, reading books, watching DVDs, and establishing other sources of support/contact. Each time we ended a session, she committed herself to beginning sobriety that day, but the pattern of drinking continued over a 5-week period. She admitted being reluctant to attend AA, feeling she needed a companion to attend at least at the outset. She identified the wine glasses in her home as another trigger for her alcohol use and agreed to pack them away. Concerned that I needed to more carefully assess the extent of her depression and better monitor its relation to her drinking, I decided to incorporate the STAR*D QIDS-SR_{16} questionnaire at week 12. She scored 14, in the moderate range, and could be considered for the next step in the STAR*D algorithm. However, given her general improvement in depression, I decided to wait before making further medication changes.

By week 13, she had incorporated some changes into her life including exercise, meditation, reading, contacting sources of support, and attending AA again. Despite these many achievements, she craved alcohol and continued to drink. She mentioned during this session that she disliked that in AA individuals usually say their name and announce "I am an alcoholic" before speaking. I wanted to probe further but instead found myself highlighting her accomplishments. I made a note to myself to address this next week. We agreed she would make a list of rewards to use instead of alcohol and to continue attending AA.

Visit 12 (Week 14)

The patient was abstinent for 1 week, an achievement she was clearly proud of. She continued to attend AA and to exercise. Although I recognized her achievements, I felt I needed to address the lingering issue of whether she viewed alcohol as a problem in her life. I began by addressing her hesitancy to say "I am an alcoholic" in AA. She first gave reasons for why alcohol might not be a problem for her but in the process realized her use of alcohol was not as controlled as it had been in the past. She then wanted to know if I thought she was an alcoholic. I responded first recognizing her accomplishments, then moved toward recognizing her struggles in trying to abstain, careful to avoid argumentation and expressing empathy throughout. In line with the principles of motivational interviewing, I also noted discrepancies between her continued use of alcohol and the goals she set for herself each week to stop (Table 34–4). Ultimately, I did say that I believed she was an alcoholic. She was saddened, realizing she could never drink again if this was true, and seemed regretful that she had been lenient with herself in returning to alcohol. We discussed other barriers to her viewing alcohol as a problem, including family who viewed alcohol favorably.

TABLE 34–4. Five general principles of motivational interviewing

Principle	Explanation	Followed
1. Avoid arguments	Counselor avoids arguing with clients. Counselor does not try to "break through" the denial, but works around it. Clients are not given diagnostic labels such as "alcoholic" or "addict." Persuasion is gentle and subtle.	√
2. Express empathy	Counselor sees the world through the client's eyes, thinking about things as the client thinks about them, feeling things as the client feels them, sharing in the client's experiences.	√
3. Support self-efficacy	Counselor focuses efforts on helping clients stay motivated by supporting clients' sense that change is possible.	√
4. Roll with resistance	Counselor does not fight client resistance, but "rolls with it" to further explore clients' views. Clients are encouraged to develop their own solutions to the problems that they themselves have defined.	√
5. Develop discrepancy	Counselor helps clients examine the discrepancies between their current behavior and future goals. When clients perceive that their current behaviors are not leading toward some important future goal, they become more motivated to make important life changes.	√

Source. Adapted from U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment: TIP 35: Enhancing motivation for change in substance abuse treatment. There are five general principles behind motivational interviewing. 2000. Available at: http://www.dawnfarm.org/pdf/MI_Principles.pdf. Accessed June 1, 2009.

I videotaped this session, which I played for my supervisor. This session, utilizing the five general principles of motivational interviewing (Table 34–4), propelled the patient to view her alcohol use in a different light. It made her step back and question how she had been handling her use. It also allowed her to contemplate the implications of being abstinent and how not drinking would influence various aspects of her life.

Visit 13 (Week 17)

Three weeks passed until my next session with Ms. M.T. She had been out of town visiting family. Despite a spike in depression attributed largely to stressors on her return home, her alcohol use had diminished; she had abstained in the week following our last session, then had one glass of wine on her return home. She had not been attending AA but was making use of relaxation techniques. Although I continued to encourage her to attend AA, what I had failed to do in this and previous sessions was explore why she was not going to AA in the first place. I asked her how she felt about our last session, and although she was ambivalent about accepting alcohol as a significant problem in her life, her cravings "disappeared" for a time after that session. I wondered, if she had not been out of town and if we had followed up sooner on our last session, would she have been more accepting of the things we discussed?

Visit 14 (Week 19)

Two weeks passed until my next session with Ms. M.T. because of my being away at a conference. In my absence, the patient's depression peaked to its highest, with a continuation of her prior stressors involving family strife and feelings of loneliness. For the first time since I had known her, she expressed death wishes, without suicidal intent or plan but more with a wish to release the emotional pain. I recognized the vicious cycle she was in of feeling depressed and drinking as a consequence to numb the pain. Although I had been careful to monitor her mood weekly, I realized that I may not have been addressing it sufficiently given my emphasis on her alcohol use. Ms. M.T. was not interested in commitments to abstain at this point, but she did feel that any changes in her alcohol use would be difficult to achieve unless her mood was better. I reviewed her antidepressant medications, which was a reasonable combination of two antidepressants at good doses. Although I considered increasing her bupropion, I decided to hold off making pharmacological changes,

feeling her mood seemed most affected by current interpersonal stressors, which we spent the session processing.

Visits 15–16 (Weeks 20–21)

Patient's mood gradually improved over the next couple weeks, utilizing assertiveness skills, which worked well in her interpersonal relationships. She also began to exercise regularly, which she found enormously helpful. Despite significant improvements in her mood, her drinking persisted, increasing to the point of drinking daily.

I decided to reassess where she wanted to go with her alcohol use. I had reminded Ms. M.T. I would be leaving the clinic in about a month. In probing, she revealed that knowing that I would be leaving the clinic soon left her feeling unmotivated to quit alcohol. She admitted she was giving herself the allowance to drink. She felt an obligation toward me, regularly coming to the clinic despite the financial and scheduling burdens. And part of this obligation involved not wanting to disappoint me if she set a commitment to quit and failed. I felt another key moment in therapy where change might be possible, similar to our session when I addressed her awareness of alcohol as a problem. After sharing these feelings of needing to please me, she felt a weight lifted from her shoulders and readily set a commitment to stop drinking for a week and to attend AA.

This session encouraged me to look deeper into the process of psychotherapy termination. Although I had ensured that Ms. M.T. was fully informed from the beginning that my time in this resident clinic would be limited to 6 months, it was clear that my upcoming departure was influencing her life decisions in ways opposite to what we were trying to achieve. Not only did she verbally reflect this, but I also noticed spikes in her depression and drinking in the weeks when either she or I was away and unable to meet (Figures 34–1 and 34–2). I found that, in some ways, she was using my departure as a way to explain her continued drinking. I could find no evidence-based guidelines on psychotherapy termination, but I did find an article in the *Journal of Clinical Psychology* (Vasquez et al. 2008) via PubMed that discussed the consequences of inappropriate termination, ethical and clinical responsibilities of the therapist, and practice recommendations (Table 34–5). In this case, there was a strong therapeutic alliance between us, and we agreed that the process was not finished but that I would not be available to continue in her care. The article provided a similar example, and stated that regression could be an unfortunate consequence. Ms. M.T. was toying with the idea of discontinuing therapy after I left, and after my research, I felt even stronger about encouraging her to continue. I made a note to myself to address this in the next session.

Visit 17 (Week 22)

The patient's mood remained stable, but despite her intended commitment to abstain, she had used alcohol every day in the past week and had not attended AA. Unanticipated celebrations and boredom led to her drinking, although she admitted to feelings of disappointment. We discussed her boyfriend's impending move and her concern that he could influence her if he continued drinking. She set goals to discuss this with her boyfriend and to attend AA. But the most significant goal she set for herself was quitting alcohol for "an extended period," recognizing that setting a commitment to quit for a limited period was not working. It was a movement toward closing the door of leniency she had given herself for years.

During this session, we continued with pretermination counseling. I wanted to readdress her plans after I left the clinic, concerned that she might decide to halt therapy and regress. My strategy was to use the content of this session and emphasize the many changes she was about to confront in the coming months and ask her how she felt about handling it on her own without a therapist. She admitted that it could be challenging and seemed to recognize the potential pitfalls of discontinuing. We discussed her reservations, namely that therapy was work on her end and that this termination was a chance to take a break before signing up with another therapist. I agreed with her that therapy was work but then noted that in the work she had done in this clinic, she had made tremendous progress. She agreed that what might be better was a therapist closer to her home, covered by her insurance, and one where continuity could be better ensured. The plan seemed to lift some of the burdens she had been feeling and removed the ambivalence she had with continuing in therapy.

Visit 18 (Week 23)

The patient's depression continued to steadily decline, her STAR*D QIDS-SR_{16} score at 10. Accord-

TABLE 34–5. Practice recommendations for psychotherapy termination

Recommendation	Followed
Provide patients with a complete description of the therapeutic process, including termination; obtain informed consent for this process at the beginning of treatment and provide reminders throughout treatment.	√
Ensure that the psychotherapist and client agree collaboratively on the goals for psychotherapy and the ending of psychotherapy.	√
Provide periodic progress updates that include discussions of termination and, toward the end of psychotherapy, provide pretermination counseling.	√
Offer a contract that provides patients with a plan in case the psychotherapist is suddenly unavailable (including death, or financial, employment, or insurance complications).	No
Help clients develop health and referral plans for posttermination life.	√
Make sure you understand termination, abandonment, and their potential effects on patients.	
Consider developing (and updating) your professional will to proactively address unexpected termination and abandonment, including the name(s) of colleagues who will contact current patients in the case of your sudden disability or death.	No
Contact clients who prematurely terminate via telephone or letters to express your concern and offer to assist them.	NA
Use the American Psychological Association Ethics Code (2002), your state practice regulations, and consultations with knowledgeable colleagues to help guide your understanding and behavior in regard to psychotherapy termination.	√
Review other ethics codes for discussions of abandonment. The American Counseling Association (http://www.counseling.org/Resources/CodeOfEthics/TP/Home/CT2.aspx) and the American Mental Health Counselors Association (http://www.amhca.org/code) contain prohibitions against abandonment.	√
Make the topic of termination a part of your regular continuing education or professional development.	√
Be vigilant in monitoring your clinical effectiveness and personal distress. Psychotherapists who self-monitor and practice effective self-care are less likely to have inappropriate terminations or clients who feel abandoned.	√

Note. NA=not applicable.
Source. Adapted from Vasquez et al. 2008.

ing to the QIDS-SR_{16} scoring criteria (University of Pittsburgh Epidemiology Data Center 2009), this was an improvement from moderate depression at week 12 to mild depression. Her alcohol use decreased substantially as well—she had 2 drinking days in the past week, compared with her daily drinking during the previous couple of weeks. She also made courageous steps toward asking her boyfriend to not drink upon his move because of its likely influence on her. I recognized how her achievements seemed to make her feel positive, careful that my own commendations or expressions of pride could hasten any obligations she might feel toward me. This was a recommendation offered by other therapists when I presented Ms. M.T. during a case conference. I decided to readdress how she viewed her alcohol use, and although she felt she was not dependent on alcohol, she believed there was a stronger "potential" for this and firmly stated she was not going to wait for it to get out of hand.

We continued with pretermination counseling, Ms. M.T. reflecting on and proud of the progress she had made in the last year. I validated her growth and strengths. She openly shared that transitioning to a new therapist might be difficult given how much she had revealed and how hard that had been for her. I recognized her feelings and was forthcoming that the termination would be difficult for me as well, ac-

knowledging our strong therapeutic alliance. We again touched on her plans to find a therapist closer to her home to continue with after my departure. I offered to look over the list of providers covered under her insurance plan with my supervisor and provide her with recommendations. In inquiring what her future goals in therapy would be, she stated "Alcohol use, going to AA, and making sure I don't get off track"—a positive addition to her prior goals, which had primarily focused on addressing her mood.

Summary of Guideline Use

At the end of 23 weeks, I reviewed my use of the guidelines. I felt I had followed the APA *Practice Guideline for the Treatment of Patients With Substance Use Disorders* (Tables 34–1 and 34–3) relatively well and consistently. The few areas I did not carry out involved intensifying relapse prevention either through encouraging monitoring programs or obtaining more objective evidence of relapse. Both these approaches may have given the patient more accountability for her actions. Furthermore, my supervisor and I had considered medications for her cravings, but I did not pursue this, and I may have assumed prematurely she would have decided against this since she believed she had self-control of her alcohol use. A thorough discussion regarding her cravings and medications as an option could have offered her an alternate treatment route to consider.

The clinic charts (Figures 34–1 through 34–3) were excellent tools that allowed Ms. M.T. and me to view the trends of her mood as it related to her alcohol use and cravings. I would have liked to use the STAR*D assessments and guidelines more carefully and from the beginning of treatment.

Although I had had a few lectures in the course of residency on motivational interviewing, I plunged into the practice for the first time after Ms. M.T.'s initial lapse. Only later did I review the five general principles (see Table 34–4), which I felt I had covered sufficiently. The one recommendation I failed to follow was not to give clients a label such as "alcoholic." In answering Ms. M.T.'s question, I did state I believed she was an "alcoholic" which triggered her into viewing her use differently. Although I felt this was a pivotal and significant point in therapy, perhaps the emphasis could have been less on the label and more on alcohol as a problem.

I also reviewed my use of the practice recommendations for psychotherapy termination, which I was satisfied I had followed. Despite following published guidelines and recommendations, this case demonstrated and affirmed for me the need to expand one's use of resources in caring for patients (i.e., conferences, supervision, articles, etc.), especially when no further guidelines exist.

The prevailing question I had for myself at the end of my time with Ms. M.T. was how successful I had been treating her alcohol use. What measure could I use as a gauge? In the COMBINE study, about 1,300 patients were enrolled in a study that compared nine different alcohol abuse treatments (Anton et al. 2006). In this study, the main outcomes were percentage of days abstinent over the 16 weeks of the study and time to first heavy drinking (five standard drinks per day for men, four for women). The mean percentage of days abstinent was about 65%, or 2 days out of 3, which was considered a minimal standard. Using this study's criteria, my patient did indeed achieve a level of abstinence above the minimal standard, at 68%.

Ways to Improve Practice

In reviewing my practice, although I had encouraged family therapy, I might have improved my care of Ms. M.T. by inviting her boyfriend and perhaps her son to one of our sessions because there were interpersonal stressors. This would have allowed me to gather collateral information, explore interpersonal dynamics, examine what understanding her family had of her drinking, and perhaps discuss ways in which they could assist with her goals. Second, I would have more carefully monitored and inquired about how our missed sessions were affecting her. Had I noticed this pattern earlier, I could have been more vigilant about checking in via telephone if we were unable to meet in person. Third, I might have been more aggressive with following the STAR*D algorithm. Her QIDS-SR_{16} was still above the optimal level for remission at our last visit. Lastly, a consultation with experts in the field of substance use/addiction could have offered further guidance in helping my patient achieve abstinence for a sustained period.

References

American Psychiatric Association: American Psychiatric Association Practice Guidelines for the Treatment of Psychiatric Disorders 2004. Treating Substance Use Disorders: A Quick Reference Guide. 2006. Available at: http://www.psychiatryonline.com/pracGuide/pracGuideTopic_5.aspx. Accessed June 1, 2009

American Psychiatric Association: Practice Guideline for the Treatment of Patients With Substance Use Disorders, 2nd Edition. 2006. Available at: http://www.psychiatryonline.com/pracGuide/loadGuidelinePdf.aspx?file=SUD2ePG_04-28-06. Accessed June 1, 2009

American Psychological Association: Ethical principles of psychologists and code of conduct. Am Psychol 57:1060–1073, 2002

Anton RF, O'Malley SS, Ciraulo DA, et al: Combined pharmacotherapies and behavioral interventions for alcohol dependence. JAMA 295:2003–2017, 2006

Center for Substance Abuse Treatment: Substance abuse: clinical issues in intensive outpatient treatment. Rockville, MD, Subxstance Abuse and Mental Health Services Administration, 2006. Available at: http://www.guideline.gov. Accessed June 1, 2009

Hunt T: Prevent a lapse becoming a relapse. Bright Eye: Online Alcoholism and Addiction Counseling. 2007. Available at: http://www.brighteyecounselling.co.uk/alcohol-drugs/prevent-a-lapse-becoming-a-relapse. Accessed June 1, 2009

Michigan Quality Improvement Consortium: Screening, diagnosis and referral for substance use disorders. Southfield, MI, Michigan Quality Improvement Consortium, 2007. Available at: http://www.guideline.gov. Accessed June 1, 2009

Sadock BJ, Sadock VA: Kaplan and Sadock's Synopsis of Psychiatry, 9th Edition. Philadelphia, PA, Lippincott Williams & Wilkins, 2003, p 412

Scottish Intercollegiate Guidelines Network: The management of harmful drinking and alcohol dependence in primary care: a national clinical guideline. Edinburgh, Scottish Intercollegiate Guidelines Network, 2003. Available at: http://www.guideline.gov. Accessed June 1, 2009

University of Pittsburgh Epidemiology Data Center: Inventory of Depressive Symptomatology (IDS) and Quick Inventory of Depressive Symptomatology (QIDS)—Scoring. 2009. Available at: http://www.ids-qids.org/index2.html#SCORING. Accessed June 1, 2009

U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment: TIP 35: Enhancing motivation for change in substance abuse treatment. There are five general principles behind motivational interviewing. 1999. Available at: http://www.dawnfarm.org/pdf/MI_Principles.pdf. Accessed June 1, 2009

Vasquez MJ, Bingham RP, Barnett JE: Psychotherapy termination: clinical and ethical responsibilities. J Clin Psychol 65:653–665, 2008

APPENDIX A

Glossary

absolute risk increase (ARI) Difference between experimental event rate (EER) and control event rate (CER) when a treatment increases the risk of a negative outcome.

absolute risk reduction (ARR) Difference between control event rate (CER) and experimental event rate (EER) when a treatment decreases the risk of a negative outcome.

allocation concealment Refers to whether the person making the allocation of a patient to either the experimental or control treatment in a randomized controlled trial is aware of the group to which the next patient will be assigned.

alpha (α) Probability of a type I error (i.e., falsely concluding that there is a difference between the experimental and control groups when such a difference is the result of chance alone).

attributable risk Difference in incidence rates between exposed and nonexposed cohorts; also known as *risk difference* (RD).

beta (β) Probability of a type II error (i.e., falsely concluding that there is no difference between the experimental and control groups when such a difference in fact exists).

bias Systematic deviation of the results from the truth.

blinded or blind Refers to whether patients, clinicians, raters, and data analysts are aware of which treatment a patient in a trial is receiving. Although terms like *single blind*, *double blind*, or *triple blind* are sometimes used, there is no consistent meaning to them, and it is better to specify which participants have been blinded.

case report Description of a single case.

case series Description of a series of cases.

case-control study Observational study in which exposure to a suspected risk factor is assessed in cases with the disease and in control subjects without the disease.

cohort A group of persons followed over time.

cohort study An observational study in which a cohort is followed over time and the number of cases of disease (or another outcome measure) is assessed. Typically, the cohort is divided into those who are exposed to a potential risk factor and those who are not exposed, and the difference or ratio of incidence rates is computed.

co-intervention Intervention other than the treatment under study that is applied differentially to the control and experimental groups.

confidence interval (CI) Range of values within which the true value most likely lies.

confounding variable Variable related to the outcome, which differs in frequency between the experimental (or exposed) and control (or nonexposed) groups. Confounding variables, not the variable under study, may be responsible for observed differences in the groups.

control event rate (CER) Rate of events in the control (nonexposed or nonexperimental) group. Typically expressed in terms of negative events.

cross-product ratio Odds ratio.
cross-sectional survey Observational study design in which exposures and outcomes are determined at the same point in time.

dichotomous outcome Outcome that can take only two values (e.g., dead or alive).

effect size Measure of the difference in outcomes between the control and experimental groups.
experimental event rate (EER) Rate of events in the experimental (exposed) group. Typically expressed in terms of negative events.

Hawthorne effect Performance improves when subjects know they are being studied.

inception cohort Group of people assembled at a common point, early in the course of their illnesses.
incidence Number of new cases of disease in a population in a given period of time.
intention-to-treat analysis Statistical analysis that includes all patients assigned to a treatment group, regardless of whether they completed study.

Kaplan-Meier curve Survival curve.

likelihood ratio (LR) Relative likelihood of test results in those with disease, compared with those without disease.

meta-analysis Statistical technique that pools results from more than one study to yield a summary result.

negative predictive value (NPV) Proportion of negative test results that are true negatives.
number needed to harm (NNH) Number of patients that must be treated with an experimental therapy to produce one additional bad outcome that would not have occurred on the control therapy. Inverse of *absolute risk increase*.
number needed to treat (NNT) Number of patients that must be treated with an experimental therapy to prevent one bad outcome that would have occurred on the control therapy. Inverse of *absolute risk reduction*.

odds Ratio of the probability of an event occurring to the probability of an event not occurring.
odds ratio (OR) Ratio of odds of event in exposed group to odds of event in nonexposed group. In a case-control study, ratio of odds of exposure in cases to odds of exposure in control subjects; also known as *cross-product ratio*.

patient expected event rate Expected risk of negative outcomes in patient if control treatment is administered.
positive predictive value (PPV) Proportion of positive test results that are true positives.
posttest odds Odds of disease in patient after results of test are known.
posttest probability Probability of disease in patient after results of test are known.
pretest odds Odds of disease in patient before results of test are known.
pretest probability Probability of disease in patient before results of test are known.
prevalence Proportion of persons in a population with disease of interest.

randomization Allocation of subjects to groups by chance.
randomized controlled trial (RCT) Clinical trial in which patients are randomly allocated to control or experimental treatment groups.
receiver operating characteristic curve (ROC) Plot of false positive rate (1–specificity) versus true positive rate (sensitivity).
relative risk (RR) In a cohort study, ratio of incidence of disease in exposed group to incidence in nonexposed group. In a randomized controlled trial, ratio of rate of negative events in experimental group (EER) to rate in control group (CER).
relative risk reduction (RRR) In an randomized controlled trial, proportion of risk of negative outcome reduced by the experimental therapy. Calculated as 1–relative risk (RR).
reliability Degree to which results are reproducible.
risk difference (RD) Difference in incidence rates between exposed and nonexposed cohorts.
risk factor Patient characteristic that increases risk of disease.
risk ratio Relative risk.

sensitivity Proportion of persons with a disease who are correctly identified by a diagnostic test.

specificity Proportion of persons without a disease who are correctly identified by a diagnostic test.

survival analysis Method of analyzing time-to-event data.

survival curve Graphical display of proportion of individuals who have not had event occur, plotted over time.

systematic review Literature review that involves comprehensive search for relevant articles, followed by critical appraisal and summarizing of articles that meet quality criteria.

type I error Falsely concluding that there is a difference between experimental and control groups when such a difference is the result of chance alone.

type II error Falsely concluding that there is no difference between experimental and control groups when such a difference in fact exists.

validity Degree to which results of a study are unbiased.

This page intentionally left blank

APPENDIX B

Statistical Formulas and Tables

B–1: Confidence intervals (CIs) for number needed to treat and number needed to harm

B–2: Calculating patient-specific number needed to treat and number needed to harm estimates

B–3: Calculating the likelihood of being helped versus harmed by a therapy

B–4: Calculating number needed to treat and number needed to harm from the odds ratio

B–5: Interpreting standardized effect size

Bibliography

Altman DG: Confidence intervals for the number needed to treat. BMJ 317:1309–1312, 1998

Freemantle N, Geddes J: Understanding and interpreting systematic reviews and meta-analyses, II: meta-analyses. Evid Based Ment Health 1:102–104, 1998

Leung R: Appendix: calculations, in Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Edited by Guyatt GH, Rennie D. Chicago, IL, AMA Press, 2002, pp 659–664

Sackett DL, Straus SE, Richardson WS, et al: Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition. New York, Churchill Livingstone, 2000

TABLE B–1. Confidence intervals (CIs) for number needed to treat and number needed to harm

	Experimental treatment	Control treatment
Bad outcome	A	B
Good outcome	C	D
Total	n_1	n_2

Control event rate (CER)=B/n_2

Experimental event rate (EER)=A/n_1

If treatment leads to fewer bad outcomes:

Absolute risk reduction (ARR)=CER–EER

Number needed to treat (NNT)=1/ARR

Standard error of ARR (SE_{ARR})= $\{[(CER)(1–CER)/n_2]+[(EER)(1–EER)/n_1]\}^{1/2}$

Upper limit of 95% CI for ARR (U_{ARR})=ARR+1.96 SE_{ARR}

Lower limit of 95% CI for ARR (L_{ARR})=ARR–1.96 SE_{ARR}

Upper limit of 95% CI for NNT (U_{NNT})=1/L_{ARR}

Lower limit of 95% CI for NNT (L_{NNT})=1/U_{ARR}

If treatment leads to more bad outcomes:

Absolute risk increase (ARI)=EER–CER

Number needed to harm (NNH)=1/ARI

Standard error of ARI (SE_{ARI})= $\{[(CER)(1–CER)/n_2]+[(EER)(1–EER)/n_1]\}^{1/2}$

Upper limit of 95% CI for ARI (U_{ARI})=ARI+1.96 SE_{ARI}

Lower limit of 95% CI for ARI (L_{ARI})=ARI–1.96 SE_{ARI}

Upper limit of 95% CI for NNH (U_{NNH})=1/L_{ARI}

Lower limit of 95% CI for NNH (L_{NNH})=1/U_{ARI}

TABLE B–2. Calculating patient-specific number needed to treat and number needed to harm estimates

	Experimental treatment	Control treatment
Bad outcome	A	B
Good outcome	C	D
Total	n_1	n_2

Control event rate (CER)=B/n_2

Experimental event rate (EER)=A/n_1

Relative risk (RR)= EER/CER

Patient's expected event rate (PEER)=estimate of patient's risk of bad outcome on control treatment

If treatment leads to fewer bad outcomes:

Relative risk reduction (RRR)=1–RR

Patient's expected absolute risk reduction (PARR)=PEER×RRR

Patient-specific number needed to treat (PNNT)=1/PARR

If treatment leads to more bad outcomes:

Relative risk increase (RRI)=RR–1

Patient's absolute risk increase (PARI)=PEER×RRI

Patient-specific number needed to harm (PNNH)=1/PARI

TABLE B–3. Calculating the likelihood of being helped versus harmed by a therapy

Step 1

Calculate the patient-specific number needed to treat (PNNT) for the desired therapeutic effect and the patient-specific number needed to harm (PNNH) for the side effect of concern (Table B-2).

Step 2

Determine the patient's relative preference (RP) for the desired effect versus the side effect. For example, if a patient feels that it is twice as bad to gain 10 kg as it is to relapse, RP=1/2.

Step 3

Calculate the likelihood of being helped or harmed (LHH), using this formula:

$$\text{LHH} = \text{PNNH}/(\text{PNNT} \times \text{RP})$$

The LHH is the relative likelihood of being helped versus harmed by a therapy, taking into account the patient's individualized risks and values.

TABLE B–4. Calculating number needed to treat and number needed to harm from the odds ratio

If treatment leads to fewer bad outcomes:

$$\text{NNT} = \frac{1 - (\text{CER})(1 - \text{OR})}{(\text{CER})(1 - \text{CER})(1 - \text{OR})}$$

If treatment leads to more bad outcomes:

$$\text{NNH} = \frac{1 + (\text{CER})(1 - \text{OR})}{(\text{CER})(1 - \text{CER})(1 - \text{OR})}$$

Note. NNT=number needed to treat; CER=estimated control event rate; OR=odds ratio; NNH=number needed to harm.

TABLE B–5. Interpreting standardized effect size

Standardized effect size (*d*) is a measure of the degree of overlap between experimental and control groups when there is a continuous outcome measure. To interpret *d*, use the chart below.

$d>0$: The average (mean) response in the experimental group is greater than this percentage of responses in the control group.

$d<0$: The average (mean) response in the experimental group is less than this percentage of responses in the control group.

d	Percentile
0	50
0.2	58
0.4	66
0.6	73
0.8	79
1.0	84
1.2	88
1.4	92
1.6	95
1.8	96
2.0	98
2.3	99

This page intentionally left blank

Index

Page numbers printed in **boldface** *type refer to tables or figures*