In This Issue:

In This Issue

Perspectives

Editorials:

Robert M.A. Hirschfeld
Are Depression and Bipolar Disorder the Same Illness?

Reviews and Overviews:

Kenneth S. Kendler
"A Gene for...": The Nature of Gene Action in Psychiatric Disorders

Images in Neuroscience:

Michael Vogel
The Cerebellum

Introspections:

Jeanne Steiner
In Sickness and Health
Images in Psychiatry:

Gregg E. Gorton

Milton Hyland Erickson, 1901–1980

New Research

Articles:

Jean A. Frazier, Sufen Chiu, Janis L. Breeze, Nikos Makris, Nicholas Lange, David N. Kennedy, Martha R. Herbert, Eileen K. Bent, Vamsi K. Koneru, Megan E. Dieterich, Steven M. Hodge, Scott L. Rauch, P. Ellen Grant, Bruce M. Cohen, Larry J. Seidman, Verne S. Caviness, and Joseph Biederman

Structural Brain Magnetic Resonance Imaging of Limbic and Thalamic Volumes in Pediatric Bipolar Disorder

Maria E. Fisfalen, Thomas G. Schulze, J. Raymond DePaulo, Jr., Leslie J. DeGroot, Judith A. Badner, and Francis J. McMahon

Familial Variation in Episode Frequency in Bipolar Affective Disorder

Ralph W. Kupka, David A. Luckenbaugh, Robert M. Post, Trisha Suppes, Lori L. Altshuler, Paul E. Keck, Jr., Mark A. Frye, Kirk D. Denicoff, Heinz Grunze, Gabriele S. Leverich, Susan L. McElroy, Jörg Walden, and Willem A. Nolen

Comparison of Rapid-Cycling and Non-Rapid-Cycling Bipolar Disorder Based on Prospective Mood Ratings in 539 Outpatients


Olanzapine Versus Lithium in the Maintenance Treatment of Bipolar Disorder: A 12-Month, Randomized, Double-Blind, Controlled Clinical Trial

Ramin Mojtabai, Daniel Herman, Ezra S. Susser, Nancy Sohler, Thomas J. Craig, Janet Lavelle, and Evelyn J. Bromet

Service Use and Outcomes of First-Admission Patients With Psychotic Disorders in the Suffolk County Mental Health Project

Glenn N. Saxe, Frederick Stoddard, Erin Hall, Neharika Chawla, Carlos Lopez, Robert Sheridan, Daniel King, Lynda King, and Rachel Yehuda

Pathways to PTSD, Part I: Children With Burns
Julie B. Kaplow, Kenneth A. Dodge, Lisa Amaya-Jackson, and Glenn N. Saxe
Pathways to PTSD, Part II: Sexually Abused Children

Miles McFall, Andrew J. Saxon, Charles E. Thompson, Dan Yoshimoto, Carol Malte,
Kristy Straits-Troster, Evan Kanter, Xiao-Hua Andrew Zhou, Cynthia M. Dougherty, and
Bonnie Steele
Improving the Rates of Quitting Smoking for Veterans With Posttraumatic
Stress Disorder

Axel Perkonigg, Hildegard Pfister, Murray B. Stein, Michael Höfler, Roselind Lieb,
Andreas Maercker, and Hans-Ulrich Wittchen
Longitudinal Course of Posttraumatic Stress Disorder and Posttraumatic
Stress Disorder Symptoms in a Community Sample of Adolescents and
Young Adults

Mark Olfson, Marc J. Gameroff, Steven C. Marcus, Ted Greenberg, and David Shaffer
National Trends in Hospitalization of Youth With Intentional Self-Inflicted
Injuries

Ann F. Garland, Anna S. Lau, May Yeh, Kristen M. McCabe, Richard L. Hough, and
John A. Landsverk
Racial and Ethnic Differences in Utilization of Mental Health Services
Among High-Risk Youths

Susan S.F. Gau, M.Y. Chong, Tony H.H. Chen, and Andrew T.A. Cheng
A 3-Year Panel Study of Mental Disorders Among Adolescents in Taiwan

Joseph R. Calabrese, Paul E. Keck, Jr., Wayne Macfadden, Margaret Minkwitz, Terence
A. Ketter, Richard H. Weisler, Andrew J. Cutler, Robin McCoy, Ellis Wilson, Jamie
Mullen, and The BOLDER Study Group
A Randomized, Double-Blind, Placebo-Controlled Trial of Quetiapine in the
Treatment of Bipolar I or II Depression

Research Units on Pediatric Psychopharmacology Autism Network
Risperidone Treatment of Autistic Disorder: Longer-Term Benefits and
Blinded Discontinuation After 6 Months

Brief Reports:

Mark Zimmerman, Iwona Chelminski, and Michael A. Posternak
Generalizability of Antidepressant Efficacy Trials: Differences Between
Depressed Psychiatric Outpatients Who Would or Would Not Qualify for an
Efficacy Trial
Patrik K.E. Magnusson, David Gunnell, Per Tynelius, George Davey Smith, and Finn Rasmussen

**Strong Inverse Association Between Height and Suicide in a Large Cohort of Swedish Men: Evidence of Early Life Origins of Suicidal Behavior?**

Alexandre Dumais, Alain D. Lesage, Aleksandra Lalovic, Monique Séguin, Michel Tousignant, Nadia Chawky, and Gustavo Turecki

**Is Violent Method of Suicide a Behavioral Marker of Lifetime Aggression?**

Justine M. Kent, Jeremy D. Coplan, Osama Mawlawi, Jose M. Martinez, Susan T. Browne, Mark Slifstein, Diana Martinez, Anissa Abi-Dargham, Marc Laruelle, and Jack M. Gorman

**Prediction of Panic Response to a Respiratory Stimulant by Reduced Orbitofrontal Cerebral Blood Flow in Panic Disorder**

Richard Meiser-Stedman, William Yule, Patrick Smith, Ed Glucksman, and Tim Dalgleish

**Acute Stress Disorder and Posttraumatic Stress Disorder in Children and Adolescents Involved in Assaults or Motor Vehicle Accidents**

---

**Communications and Updates**

**Letters to the Editor:**

SERMSAK LOLAK, ELISA DANNEMILLER, and FRANCIS ANDRES

**48,XXYY Syndrome, Mood Disorder, and Aggression**

DANIEL T. HOLMES, PHILLIP LONG, and JIRI FROHLICH

**Dysbetalipoproteinemia and Clomipramine**

ROBERT OSTROFF, MARBIELA GONZALES, and GERARD SANACORA

**Antidepressant Effect of Ketamine During ECT**

TUBA OZCANLI, BARIS UNSALVER, SAMURAY OZDEMIR, and MINE OZMEN

**Sertraline- and Mirtazapine-Induced Severe Neutropenia**

RAKESH KARMACHARYA, MARI MINO, and WILLIAM F. PIRL

**Clozapine-Induced Eosinophilic Colitis**

JEAN ADDINGTON and DONALD ADDINGTON

**Clinical Trials During the Prodromal Stage of Schizophrenia**

COLIN O’DONNELL and TIM STEPHENS

**The Significance of Homocysteine Levels in Schizophrenia**
DONALD C. GOFF

**Dr. Goff and Colleagues Reply**

ANDREW J. DWOR, GORAZD ROSOKLIJA, and LIESL B. JONES

**Reduced Spinophilin in Schizophrenia**

AMANDA J. LAW, CYNTHIA SHANNON WEICKERT, THOMAS M. HYDE, JOEL E. KLEINMAN, and PAUL J. HARRISON

**Dr. Law and Colleagues Reply**

MARK S. GOLD, KIMBERLY FROST-PINEDA, and RICHARD J. MELKER

**Physician Suicide and Drug Abuse**

EVA SCHERNHAMMER

**Dr. Schernhammer Replies**

RYAN M. CARNAHAN and PAUL J. PERRY

**Methodological Concerns in a Trial of Ziprasidone and Olanzapine**

DAVID E. ROSS

**Methodological Concerns in a Trial of Ziprasidone and Olanzapine**

GEORGE M. SIMPSON

**Dr. Simpson Replies**

LIEUWE de HAAN and NICO van BEVEREN

**Olanzapine and Haloperidol for Residual Symptoms**

ROBERT W. BUCHANAN, M. PATRICIA BALL, and ELAINE WEINER

**Dr. Buchanan and Colleagues Reply**

JACK C. SCHOENHOLTZ

**Origin of the Term "Schizophrenia"**

---

**Book Forum:**

SAXBY PRIDMORE

**The Neurological Basis of Pain**

FRANCINE M. BENES

**Neurology for the Non-Neurologist, 5th ed.**

ROBERT HOWARD

**Neuropsychiatric Assessment**
RICHARD B. MAKOVER
   Sleep Medicine in Clinical Practice

DAVID V. FORREST
   Brain Dynamics and Mental Disorders: Project for a Scientific Psychiatry

MARK H. FLEISHER
   The Mind: Its Nature and Origin

MARY V. SEEMAN
   The Philosophy of Psychiatry: A Companion

WILLIAM M. GREENBERG
   The Biopsychosocial Approach: Past, Present, Future

LAURA L. POST
   Mind, Meaning, and Mental Disorder: The Nature of Causal Explanation in
   Psychology and Psychiatry, 2nd ed.

TRUCE T. ORDOÑA
   Feeling Good: The Science of Well-Being

DONALD M. HILTY, BLYTHE A. CORBETT, and PIERRE LAVENEX
   Affect Dysregulation and Disorders of the Self

GORDON PARKER
   Experiences of Depression: Theoretical, Clinical, and Research Perspectives
Brains of Bipolar Kids
Compared to adults with bipolar disorder, children and adolescents with this illness have more severe symptoms, a stronger genetic predisposition, and less past drug treatment. Their brains might therefore provide a clearer picture of anatomic abnormalities in bipolar disorder. Frazier et al. (p. 1256) measured three brain regions in 43 youths with bipolar disorder and 20 healthy youths. The amygdala and thalamus showed no difference between groups, but the hippocampus, which provides context for learning and memory, was smaller in the bipolar youths. The bipolar girls showed an especially pronounced difference from the same-sex comparison subjects, and their mania scores were higher than those of the bipolar boys. The total cerebral volume in the bipolar group was also smaller than normal. Since total cerebral volume generally is normal in affected adults, bipolar disorder in children and adolescents may follow a different neurodevelopmental path.

The Incredible Shrinking Hospital Stay
From 1989 to 1995, the average length of hospitalization for psychotic disorders in Suffolk County, N.Y., plummeted from 41 to 21 days. Nevertheless, Mojtabai et al. (p. 1291) found that posthospitalization symptoms and illness course and outcomes were no worse among patients admitted in 1995 than among those hospitalized in 1989. There was also no significant change in the number of rehospitalizations, the rate of suicide or homicide, or the use of outpatient mental health services. In fact, patients admitted in the later years were more likely to have returned to their highest level of functioning by 4-year follow-up. Introduction of atypical antipsychotic drugs did not account for this lack of decline in patient outcomes. One troubling finding was the greater likelihood of partial or no remission at discharge among patients hospitalized later in the study period. Linkage to outpatient services is especially important for these patients.

Protecting Burned Children From PTSD
Fire is a common cause of accidental injuries to children. Little is known about which of these children develop posttraumatic stress disorder (PTSD) and which do not. Saxe et al. (p. 1299) found two separate causal pathways to PTSD in 72 children admitted to a hospital for burn victims. Separation anxiety and dissociation of mental processes (e.g., amnesia for traumatic events) independently contributed to PTSD at 3 months after hospitalization. Each was itself influenced by the size of the burn. Separation anxiety was also greater for younger children and those with greater pain. Anxiety and dissociation appear to reflect different biological reactions to overwhelming stress—a fight-or-flight response versus freezing/immobilization. These two pathways to PTSD suggest specific interventions to help burned children: increased availability of their parents in the hospital, support of the parents, and psychosocial and biological treatments for pain, anxiety, and dissociation.

Kids Hospitalized for Self-Injury
Between 1990 and 2000, the overall rate of hospitalization for children and adolescents who intentionally injured themselves declined only slightly. However, the average length of these hospital stays decreased significantly, and diagnoses shifted toward more severe conditions, e.g., bipolar disorder. Olfson et al. (p. 1328) derived these findings from over 10,000 hospitalizations for self-inflicted injuries of patients ages 5–20 years. Although the rate did not change overall, hospitalizations increased for children ages 5 to 9 years, for ingestion of acetaminophen and antidepressants, and for injuries from cutting and from hanging or suffocating. Cutting may be more closely related to self-mutilation than to suicidal behavior, however, and without it the rate of hospitalizations for self-inflicted injuries shows an actual decline, from 47.2 per 100,000 population in 1990 to 39.4 per 100,000 in 2000.

Race of High-Risk Youths Affects Use of Mental Health Services
Many children and adolescents in child welfare programs, the juvenile justice system, and other public care sectors need mental health services. The Surgeon General’s Report on Mental Health has identified racial/ethnic disparities in use of mental health services as a major public health problem. Garland et al. (p. 1336) examined whether ethnicity itself, apart from associated factors such as socioeconomic status, influenced the use of mental health services by 1,256 children and teens in five types of public care in San Diego County. Compared to non-Hispanic whites, African Americans and Asian Americans were approximately half as likely to receive outpatient services, as well as mental health services in general. Other predictors, such as caregiver strain, were also identified and suggest that attention to family- and system-level factors could open doors to young people needing help.
Are Depression and Bipolar Disorder the Same Illness?

Along a long-standing scientific controversy with many clinical consequences is whether bipolar and unipolar disorder are the same or separate and distinct illnesses. Two articles in this issue of the Journal present findings that address this controversy with new evidence.

Depression has long been viewed as a fundamental human condition. Descriptions of the syndrome of depression are included in the writings of Hippocrates from 2,500 years ago. He attributed melancholic temperament to an excess of black bile emanating from the liver, one of four bodily humors (1). The symptoms of dysphoria, psychomotor retardation, and suicidality are consistent with DSM-IV criteria.

In contrast, the concept of bipolarity as a fundamental human condition is quite new. Although “mania” was mentioned by ancient Greek physicians, including Hippocrates, its description varied widely, with little consistency (2).

In the late 19th and early 20th century, Kraepelin, in Germany, emphasized the distinction between dementia praecox and manic-depressive insanity (3, 4). One source of this dichotomy was the emphasis on moods versus cognition and will. However, the major rationale for the distinction was a perceived difference in clinical course and outcome. Dementia praecox was characterized by a deteriorating course, whereas mood disorder insanity was characterized by a relapsing course (5). His concept of manic-depressive insanity included both recurrent unipolar and bipolar illness.

It was not until 1966 that bipolar disorder was described as separate and distinct in articles by Jules Angst (6), Carlo Perris (7), and Winokur and colleagues (8) in the United States. These articles proposed separate unipolar and bipolar disorders based on difference in genetics, gender, clinical course, and premorbid personality.

The two articles in the current issue of the Journal add to the growing literature supporting the distinctness of bipolar disorder. In the article by Fisfalen and colleagues, the authors report that the frequency of illness episodes is highly familial among patients with bipolar disorder compared with recurrent unipolar depression. This finding serves to provide evidence of a genetic difference between bipolar disorder and unipolar disorder.

The article by Frazier and colleagues also addresses the issue of bipolar disorder as a distinct illness with a biological basis. The authors found structural differences in components of the limbic system in prepubertal children with bipolar disorder. Since regulation of mood and basic human drives (e.g., sleep, sex, and appetite) is a function of the limbic system, these structural differences relate to causes of the illness.

A very intriguing finding in the study by Frazier et al. is that of decreased hippocampal volumes in children with bipolar disorder, which has not been reported in adults. This raises the question of whether prepubertal bipolar disorder is somehow different from adolescent- or adult-onset bipolar disorder. Consistent with this is a difference in the presenting clinical practice of mania in children (9). Instead of a euphoric mood, the most frequent mood disturbance is severe irritability, often with accompanying protracted, hostile, and violent temper outbursts (10). During these affective storms, it is nearly impossible to calm the child. This situation tends to be persistent rather than episodic (as is seen more often in adults). At the same time, classic symptoms of mania oc-
cur in these children. These include grandiosity, hypersexuality, increased energy, and decreased need for sleep. It is intriguing to speculate whether these hippocampal imaging differences between children on one hand and adolescents and adults on the other are responsible for the differences in clinical presentation.

References
3. Kraepelin E: Psychiatrie, 4. Leipzig, Germany, JA Barth, 1893
5. Angst J: Historical aspects of the dichotomy between manic-depressive disorders and schizophrenia. Schizophr Res 2002; 57:5–13

ROBERT M.A. HIRSCHFELD, M.D.

Address correspondence and reprint requests to Dr. Hirschfeld, Department of Psychiatry and Behavioral Sciences, University of Texas Medical Branch, 301 University Blvd., Galveston, TX 77555; rohirsch@utmb.edu (e-mail).
"A Gene for...": The Nature of Gene Action in Psychiatric Disorders

Kenneth S. Kendler, M.D.

A central phrase in the new "GeneTalk" is "X is a gene for Y," in which X is a particular gene on the human genome and Y is a complex human disorder or trait. This article begins by sketching the historical origins of this phrase and the concept of the gene-phenotype relationship that underlies it. Five criteria are then proposed to evaluate the appropriateness of the "X is a gene for Y" concept: 1) strength of association, 2) specificity of relationship, 3) noncontingency of effect, 4) causal proximity of X to Y, and 5) the degree to which X is the appropriate level of explanation for Y. Evidence from psychiatric genetics is then reviewed that address each of these criteria. The concept of "a gene for..." is best understood as deriving from preformationist developmental theory in which genes—like preformationist anlagen—"code for" traits in a simple, direct, and powerful way. However, the genetic contribution to psychiatric disorders fails to meet any of the five criteria for the concept of "X is a gene for Y." The impact of individual genes on risk for psychiatric illness is small, often nonspecific, and embedded in complex causal pathways. The phrase "a gene for..." and the preformationist concept of gene action that underlies it are inappropriate for psychiatric disorders.

The last 20 years has seen the rise of "GeneTalk" (1). A central phrase in GeneTalk, and one that has been heard widely in both lay (2) and professional arenas, is "X is a gene for Y," in which X is a particular gene on the human genome and Y is one of a wide variety of complex human disorders or traits such as depression, aggression, sexual orientation, obesity, infidelity, alcoholism, or schizophrenia.

This essay begins with a brief review of the historical origins of the concept of "a gene for...". I then propose criteria to assess the validity of this model of gene-phenotype relations and go on to evaluate these criteria as applied to genetic effects on psychiatric disorders. The essay concludes with general observations about our preconceptions and the reality of gene action in psychiatric disorders. Although many of the issues raised in this essay are equally applicable to etiologically complex medical disorders, the focus here will be on psychiatric illness.

Historical Origins of the Concept of "A Gene for..."

Since humans started speculating about the nature of development and inheritance, a number of different conceptualizations have emerged about the nature of the guiding forces in these processes (3). In the 20th century, this discourse has come to focus largely on the nature of what Mendel originally termed "anlagen" or "elements," which in 1909 became "genes" (4).

Of the multiple different views of the nature of the "gene," the one in which we are interested—a gene defined by the phenotype that it causes—originated in the developmental theory of preformationism (5). One of the earliest articulated theories of development, preformationism was first proposed by Aristotle but became particularly influential in the 17th century (3, 5, 6). The essentials of the theory are eloquently described by Jacob:

At a time when living beings are known by their visible structure alone, what has to be explained about generation [i.e., development] is the maintenance of this primary structure through succeeding generations. The structure cannot itself disappear; it has to persist in the seed from one generation to another. To maintain the continuity of shape, the "germ" of the little being to come has to be contained in the seed; it has to be "preformed." The germ already represents the visible structure of the future child....It is the plan for the future living body...already materialized, like a miniature of the organism to come. It is like a scale model with all the parts, pieces and details already in position....Fertilization only activates it and starts it growing. Only then can the germ develop, expand in all directions and acquire its final size, like those Japanese paper flowers which, when placed in water, unwind, unfold and assume their final shape. (7, p. 57)

In preformationism, the egg or sperm was understood to contain all the final traits of the mature organism. Development consisted of the expansion of these preformed characteristics (or anlagen) into the individual traits of the adult organism. That is, these anlagen were truly for the adult traits with which they had a simple and direct causal relationship.
In the 19th century, as the young field of biology struggled to fathom the mechanism of transmission of traits across generations, a number of the proposed theories of inheritance (where the “units” of inheritance had names such as pangenes, stirps, and gemmules) had important preformationist themes (3, 4). When Mendel’s groundbreaking work on genetics (originally published in 1866) was rediscovered in 1900, one common interpretation was that his “elements of inheritance” were the discrete anlagen predicted by preformationist theories (5). This interpretation was favored by two of the most influential geneticists of the day, the Dutchman de Vries (the most famous of the three “co-rediscoverers” of Mendel [4]) and the Englishman Bateson (8).

In summarizing this exciting period in the history of biology, Allen (9) writes

The implications that the discreteness of the gene implied the organism was constructed as a “mosaic” of adult traits was given explicit voice by Bateson with the first years of his encounter with Mendelism.

Allen goes on to quote two passages from Bateson written, respectively, in 1901 and 1902 (9):

In so far as Mendel’s law applies, the conclusion is forced upon us that the living organism is a complex of characteristics of which some, at least, are dissociable and are capable of being replicated by others. We thus reach the conception of unit characters which may be rearranged in the formation of reproductive cells.

The organism is a collection of traits. We can pull out the yellowness and plug in greenness, pull out tallness and plug in dwarfishness.

Bateson was recasting, in a new language, preformationist concepts. The Mendelian anlagen (later genes) could be defined by their relationship to the particular phenotype (or “unit character”) with which it had a privileged causal link. That is, such genes caused phenotypes in the same way that the preformationist anlagen prefigured adult traits. From this perspective, it made sense to speak of “a gene for greenness,” “a gene for tallness,” or a gene for any of the innumerable unit characteristics of the adult organism. It is in this context that a rarely discussed early chapter of psychiatric genetics in the United States must be viewed, when reports appeared claiming to find, in series of large pedigrees, evidence for Mendelian genes “for” “Nomadism or the wandering impulse” (10) and “the neuropathic constitution” (11).

This preformationist concept of the gene proved attractive to medical geneticists who, over the course of the 20th century, showed that most classical genetic disorders in humans (termed “Mendelian” diseases in honor of the Austrian monk) were due to hereditary units that behaved just like those first examined by Mendel (12).

While medical geneticists came to understand that in biological systems, genes actually code for proteins, it became convenient and seemingly natural to think about preformationist-like genes for these classical genetic diseases in humans.

The last 30 years have seen three interrelated further themes in the “a gene for...” story. First, in the mid-1970s two influential books appeared that heightened the profile of genes and their potential impact on human behavior. “Sociobiology: The New Synthesis” by Wilson (13) launched the field of sociobiology (and later evolutionary psychology), discourse in which commonly included the concept of “genes for” a wide range of traits, including altruism, territoriality, jealousy, and ethics. “The Selfish Gene” by Dawkins (14) proposed a gene-centered view of evolution in which an organism, with its wide array of phenotypes, was viewed as a vehicle through which genes replicate themselves over evolutionary time. Second, with the development of an ever increasing set of powerful molecular tools, the specific genes and then the specific mutations in those genes were discovered that were responsible for all major classic human genetic disorders. So, when speaking about “a gene for Y” in which Y was sickle-cell anemia, cystic fibrosis, or Huntington’s chorea it became possible to conceive of the gene not only as an abstract transmitted “unit” but also as a discrete piece of DNA at a specific location on a chromosome. Third, prompted by the sequencing of the human genome, the concept that DNA represented the “blueprint” of life (or in related versions the “code” or “recipe” for life) was widely promulgated in both the scientific and lay literature (2). The preformationist themes in this metaphor are evident: genes are to phenotypes as blueprints of a building are to the building themselves.

So, this historical sketch suggests that our current concept of “X is a gene for Y” in humans has four major interrelated historical roots. First, the concept that development anlagen could be “for” adult traits arose in preformationist developmental theories. Second, the discovery of Mendel’s “elements” was interpreted by some as verifying this concept. Third, the idea that genes could be “for” human traits was supported by the discovery that genes for classical Mendelian medical disorders often acted just like the hereditary elements found in Mendel’s pea plants. Finally, these concepts became linked to DNA by a series of stunning discoveries in the last 20 years, so that strength of the “icon” of the double helix provided particular luster to potential discoveries in psychiatry of “a gene for...”.

Criteria for the Concept of “A Gene for...”

The remainder of this essay addresses the question of whether this preformationist model of gene action—in which genes are “for” phenotypes—is appropriate for psy-
psychiatry. Based in part on prior efforts to develop guidelines for causal inference in epidemiology (e.g., reference 15), I suggest five criteria by which to judge the validity of the claim "X is a gene for Y": 1) strength of association of X with Y, 2) specificity of relationship of X with Y, 3) noncontingency of the effect of X on Y, 4) causal proximity of X to Y, and 5) the degree to which X is the appropriate level of explanation for Y. In sum, I argue that

If gene X has a strong, specific association with disease Y in all known environments and the physiological pathway from X to Y is short or well-understood, then it may be appropriate to speak of X as a gene for Y.

But first, a few details are needed. The scientific basis of most claims that "X is a gene for Y" results from a statistical test called association analysis. In its simplest form, this test compares the frequency of specific DNA variants in or around gene X in a set of cases with disorder Y and a set of matched control subjects. An association is claimed if the frequency of these variants differ significantly in cases and control subjects. In both a conceptual and statistical sense, this approach is no different from the methods commonly used in the biomedical and social sciences to assess the relationship between putative risk factors and outcome variables such as smoking and lung cancer or childhood sexual abuse and depression.

Therefore, standard "a gene for..." claims are based on statistical and not biological grounds. Biological studies that trace etiologic pathways from X to Y should follow claims for association and would certainly provide confirmatory data. However, they have been very rare to date in psychiatric genetics. On its own, a significant p value in an association study tells you nothing about the nature of the causal relationship between the gene and the disease.

**Strength of Association**

As with any risk factor for any outcome, the strength of association between a specific gene and a particular disease can vary in magnitude. In considering the criteria for "a gene for...", an historical standard of comparison is what has come to be called a Mendelian gene. The action of Mendelian genes is deterministic and not probabilistic. If a plant inherits a particular copy of the gene for wrinkled peas, it would not matter how much sunshine the plant received or the quality of its fertilizer. The plant will have wrinkled peas no matter what the environment does. In humans, we have many diseases that are due to Mendelian genes that behave exactly like the genes Mendel studied in his pea plants (12). If you have one copy of the pathogenic gene for Huntington's disease, it does not matter what your diet is, whether your parents were loving or harsh, or if your peer group in adolescence were boy scouts or petty criminals. If you have the mutated gene and you live long enough, you will develop the disease.

Furthermore, for most Mendelian genes in man, the only way to get the disorder is to have the disease gene. There is no way to "acquire" cystic fibrosis or Huntington's disease through environmental exposure. So if having the disease gene always produces the disorder and the disorder never occurs without the disease gene, this produces a perfect association between the disease gene (X) and the disorder (Y). (Reality is somewhat more complex. Most Mendelian genes in man contain several different mutations, each of which can cause diseases that are sometimes of quite variable severity. But this claim still holds for all mutations of the gene considered together.)

The strength of an association between a risk factor and a disease is most frequently quantified by a statistic called the odds ratio. Formally, the odds ratio is defined as the ratio of the odds of developing the disease among those exposed to the risk factor and the odds of disease among those not exposed to the risk factor. For Mendelian disorders in man, since the first of these figures is one and the second is zero, the odds ratio for the disorder given the pathogenic gene is infinite. Since this is a rather stringent criteria, for the sake of argument, let us say the association with Mendelian-like genes (an historical model for the concept of "a gene for...") has an odds ratio of approximately 100 (Figure 1).

Are there any genes whose strength of association with a psychiatric disorder is Mendelian-like? Two related sources of information, both gathered in the last two decades, indicate that the answer to this question is almost certainly "No." First, a gene that has a deterministic or nearly deterministic relationship with a phenotype produces an unmistakable signature in the pattern of illness in large pedigrees. Numerous investigators have now searched many parts of the globe (including nearly all psychiatric facilities in a modest-sized country [16]) seeking pedigrees in which major forms of psychiatric illness—especially schizophrenia and bipolar illness—are distributed in the pattern expected from a Mendelian-like gene. Such pedigrees have not been found.

Second, Mendelian-like genes also produce a distinctive result in genome-wide linkage studies, which effectively sweep the human genome looking for regions that contain genes that have an impact on risk of illness. While the technical details need not concern us, experts agree that for those disorders studied in genome-wide linkage scans of reasonable size and quality—especially schizophrenia, bipolar illness, panic disorder, and eating disorders—conclusive evidence has accumulated that even moderately rare genes of Mendelian-like effect do not exist. (The available evidence does not permit us to rule out, however, very rare Mendelian-like genes.)

So, if we lack Mendelian genes for psychiatric disorders, with their very high odds ratios, what sort of magnitude of associations might we expect? One set of benchmarks might be provided by three examples of what would be considered very strong associations in epidemiology. The estimated odds ratio between heavy smoking and small cell carcinoma of the lung is approximately 20 (17), between in-
Although the odds ratio for a classic Mendelian gene is actually 1.246, Am J Psychiatry 162:7, July 2005, “of the last decades, the association between the pathogenic outstanding genetic association results in neuropsychiatry the onset of major depression is approximately 12 (19), and between severe stressful life events and severe mesothelioma is approximately 15 (18), and between heavy smoking and lung cancer, industrial exposure to asbestos and mesothelioma, and severe stressful life events and the onset of major depression. Moderate association (odds ratio=5.0) approximates that seen for apolipoprotein E gene and Alzheimer’s disease as well as the protective effect in Asian populations of the ALDH2*2 copy of the aldehyde dehydrogenase gene on risk for alcoholism. The association seen between individual genes (or high-risk haplotypes) and psychiatric disorders (odds ratio=1.5) is an approximation obtained from a review of the current literature.

Another more modest benchmark is provided by the two outstanding genetic association results in neuropsychiatry of the last decades. The association between the pathogenic “4 allele” of the apolipoprotein E gene and Alzheimer’s disease produces, in Caucasian populations, an odds ratio of approximately 3.0 (20). In Asian populations, the possession of the slow-metabolizing (ALDH2*2) copy of the aldehyde dehydrogenase gene conveys up to a 10-fold reduction in risk for the development of alcoholism (21).

So, as depicted in Figure 1, we have three possible benchmarks for the strength of the gene-phenotype association for psychiatric disorders: Mendelian-like (odds ratio of approximately 100), strong (odds ratio=12–20), or moderate (odds ratio=3–10).

Trying to summarize the magnitude of association found between functional candidate genes and psychiatric disorders is problematic because of the multiple methodologic difficulties in the interpretation of such studies (22–24). Greatest reliability should be placed on the results of meta-analyses, which are now beginning to appear in the literature. A PubMed search from 2000 on (using publication type of “metaanalysis” and search words “gene” and “association”) followed by a hand search and elimination of duplication yielded 10 significant meta-analytic estimates of odds ratios between individual genes and psychiatric disorders (Table 1) (excluding results from those meta-analyses that did not support the original positive reports). The odds ratios ranged from 1.07 to 1.57 with a median of approximately 1.30.

Another strategy to localize candidate genes is to look for them under linkage peaks (so-called positional candidate genes). In schizophrenia, replicated evidence is now emerging for several such genes (28). For these genes, disease-associated haplotypes—small sections of DNA that have traveled together over evolutionary time—can often be found. The two best replicated positional candidate genes for schizophrenia are dysbindin 1 and neuregulin 1. Not counting the original reports (where the effect size might be biased upward), estimates are available for the association between high-risk haplotypes and schizophrenia for both of these genes. For dysbindin, odds ratios of 1.24 (29), 1.23 (30), 1.40 (31), 1.70 (32), and 1.58 (33) have been reported or calculated from replication reports. For neuregulin 1, two replications were noted in a recent review, with odds ratios estimated to be 1.25 and 1.80 (28).
tional candidate genes suggest that the magnitude of the associations between individual genes and psychiatric illnesses have small odds ratios, largely from 1.1 to 1.6. Compared to our benchmarks, this effect size is very modest (Figure 1). Perhaps genes (or particular mutations or haplotypes) of larger effect size will be found. While results from linkage studies suggest that this is unlikely, it cannot be ruled out. Also to be considered is the statistical dictum that the first set of effects detected in any research area tend to be the most robust. If this is correct, further genes discovered for psychiatric disorders are likely to have smaller average effects than the genes found to date.

The preformationist concept of “a gene for...” implied a predetermined and largely irrevocable link between gene and phenotype. This is the pattern of association observed between gene and phenotype from Mendel’s original traits and for Mendelian genetic disorders in humans. By contrast, for psychiatric disorders, individual genes appear to have a quite modest association with psychiatric illness. While they may have an impact on risk, individual genes hardly predetermine illness, as would be expected if we had discovered “genes for” mental disorders.

Specificity of Association

The second criterion to evaluate the appropriateness of the concept of “X is a gene for Y” is the degree of specificity in the relationship between X and Y. As illustrated in Figure 2, does X influence risk for any other disorders in addition to Y? Or are there other genes that contribute to Y in addition to X?

In preformationist theory, anlagen had highly specific associations with the adult traits into which they developed. The hereditary elements of the pea that Mendel studied also had quite specific phenotypic effects. That is, one gene influenced pea color but not shape or height while another influenced shape but not height or color. However, as genetics developed, many genes were found that impacted on a variety of phenotypic characteristics—a phenomenon called pleiotropy.

In man, many Mendelian genes produce one and only one disease syndrome (although sometimes of varying severity depending on the specific mutation). But there are exceptions where different abnormalities in a single gene can produce distinct genetic diseases.

How specific are individual genes in their impact on risk for psychiatric disorders? Do most genes influence risk for one and only one psychiatric disorder? Twin studies, which study “genes” in the aggregate, suggest that genetic risk factors for psychiatric disorders are often nonspecific in their effect. A large-scale twin study of seven psychiatric and substance use disorders found one common genetic risk factor predisposing to drug abuse, alcohol dependence, antisocial personality disorder, and conduct disorder and a second common genetic factor influencing risk for major depression, generalized anxiety disorder, and phobia (34). Overlap of genetic risk factors for multiple disorders have been demonstrated in other twin studies (e.g., references 35–37).

We know much less about the specificity of the spectrum of effects on psychiatric disorders of individual genes. Meta-analyses reviewed in Table 1 show that variants at one gene (the 5-HT2A receptor) may predispose to risk for three different disorders (schizophrenia, bulimia, and anorexia nervosa). A pair of overlapping genes on chromosome 13q (termed G30 and G72) may be associated both with schizophrenia and bipolar illness (28). A number of overlapping positive regions in linkage genome scans for bipolar illness and schizophrenia have led some to argue that this reflects shared genes between these two disorders (38). While difficult to evaluate critically, claims have been made that several popular candidate genes (e.g., serotonin transporter, dopamine transporter, dopamine 2 receptor) are significantly associated with a wide variety of psychiatric disorders or psychiatrically relevant traits (39, 40). While much remains unknown, current evidence suggests that many genes that influence risk for psychiatric disorders will not be diagnostically specific in their effect, thereby resembling the one-to-many relationship in Figure 2 rather than the one-to-one relationship.

We are on firmer ground in evaluating whether genetic risk for psychiatric disorders results from the action of a single gene (the one-to-one relationship in Figure 2) or multiple genes (the many-to-one relationship in Figure 2).

![FIGURE 2. Possible Gene-to-Phenotype Relationships](image-url)
Some evidence bears on this question indirectly, as follows. Twin and adoption studies provide convincing evidence for significant genetic effects on virtually all major psychiatric disorders (41). Therefore, genes that affect risk for these disorders must exist somewhere on the human genome. Linkage studies examine how these aggregate genetic risk factors are distributed across the genome. If genetic risk resulted from a single gene, then all the linkage “signal” would be concentrated in a single location, with a resulting clear and robust statistical linkage peak. But, as noted earlier, this is a pattern that has not been observed in published genome scans for psychiatric disorders. Instead, a number of modest linkage peaks are usually seen, suggesting that the “packets” of genetic risk for these disorders are widely dispersed across the genome. (To complicate matters, genome scans will underestimate the number of genomic regions involved because of low power to detect genes of small effect size, but will overestimate the number because some of the observed “peaks” will be false positives.)

Recently, data have emerged that addresses this question directly. A careful meta-analysis of 20 genome scans for schizophrenia has suggested 10 genomic regions likely to contain susceptibility genes (42). In addition, current evidence of bipolar disorder, the second-best-studied psychiatric disorder by linkage scans, also suggests multiple loci (43).

The specificity of association implied in the “a gene for…” concept has another implication worth exploring. Consistent with preformationist theory, specificity of gene action implies that the gene contains all information needed for the development of the trait. The environment might impact on the final phenotype, but its effect is non-specific. That is, the gene “codes for” the trait, while the environment reflects background factors that support development but is not in and of itself “information-carrying.”

To illustrate how commonly we see genes and environment in this light, it is worth pondering a curious and asymmetrical feature of GeneTalk. While we find it easy to use the phrase “X is a gene for Y,” it feels quite odd to say “A is an environment for B.” For example, a large body of empirical work supports the hypothesis that severe life events are important environmental risk factors for major depression (44). The magnitude of the association between such events and the subsequent depressive episode is far greater than that observed for any of the genes that we have reviewed here. Yet, who has heard the phrase “a romantic breakup is an environment for depression”? I suggest that we feel comfortable with “X is a gene for Y” and not “A is an environment for B” because we implicitly assume that genes have a privileged causal relationship with the phenotype not shared by environmental factors.

However, empirical evidence does not support the position that genes code specifically for psychiatric illness while the environment reflects nonspecific “background effects.” By definition, environmental factors are central to the etiology of posttraumatic stress disorder. In the aforementioned multivariate twin model, what distinguished major depression, generalized anxiety disorder, and phobia from one another were environmental and not genetic risk factors (34). In a detailed study of the impact of childhood parental loss on risk for common psychiatric and substance use disorders, death of a parent was specific in increasing risk for major depression and no other disorder (Kendler et al., unpublished results). Consistent with studies of stressful life events that have shown moderate separation of depressogenic and anxiogenic events (45, 46), a multivariate genetic study of symptoms of anxiety and depression showed that genetic factors influence nonspecific risk for all symptoms, whereas two environmental factors were identified that predisposed, with moderate specificity, for symptoms of depression and anxiety, respectively (47).

The preformationist concept of “a gene for…” implies high levels of specificity between gene and phenotype. While much remains to be learned in this area, current evidence suggests that instead of the “one-to-one” relationship implied by the concept of “a gene for…,” genes and disorders in psychiatry are likely to have the “many-to-many” relationship depicted in Figure 2.

(The evidence that the association between individual genes and psychiatric disorders are typically weak and may often be nonspecific does not mean that the identification of such genes is unimportant. For example, such discoveries can identify pathophysiologic pathways, begin the lengthy process of clarifying how individual genes interact with each other and with environmental exposures to produce illness, and provide new targets for treatment.)

Noncontingency of Association

Noncontingent association means that the relationship between gene X and disorder Y is not dependent on other factors, particularly exposure to a specific environment or on the presence of other genes. As mentioned earlier, this is a typical (albeit not uniform) feature of genes that cause classical Mendelian disorders in humans. If the association between gene and disease were contingent on particular environmental exposures, then we would have to amend our statement to read “X is a gene for Y given exposure to environment Z.”

Environmental contingencies for genetic effects on psychiatric disorders have been little investigated. Twin and adoption studies suggest that the impact of aggregate “genes” for major depression are altered by exposure to stressful life events (19, 48) and for schizophrenia and conduct disorder by exposure to a dysfunctional rearing environment (49, 50). A range of twin studies suggest that environmental experiences have an impact on genetic risk for several psychiatrically relevant traits, including aggression, disinhibition, and smoking (51). Recently, Caspi and colleagues have found evidence for interactions between environmental risk factors and particular genes in the pro-
duction of antisocial behavior (52) and depression (53), with the former finding having been replicated (54).

We know almost nothing about gene-by-gene interactions in the etiology of psychiatric disorders. Although a number of association studies have reported interactions, I am unaware of any that have been widely replicated or supported by meta-analyses. Using statistical models applied to risk of illness in various classes of relatives, Risch has claimed that gene-by-gene interactions are important in the etiology of schizophrenia (55).

Overall, we know little about the contingent nature of genetic effects for psychiatric disorders. The available information suggests that gene action contingent upon certain environmental exposures is probably not rare and may be relatively common for psychiatric disorders. This is also inconsistent with the preformationist concept of “a gene for...”.

**Causal Proximity**

Preformationist developmental models assumed that anlagen developed directly into adult traits. The “blueprint for life” metaphor similarly assumes a direct correspondence between individual parts of the blueprint (windows, doors, fixtures) and the corresponding units of the completed building. Conceptualizing genes in this preformationist framework therefore carries the implicit assumption of a direct causal link between gene and phenotype. It is only with this assumption that usage of the “a gene for...” is congruent with the common meaning of the phrase “X is for Y” in English. To clarify this point, let’s examine a typical list of such statements:

- I use a knife for buttering my toast.
- I have a backpack for carrying my computer to work each day.
- I was upset at my son for not doing his chores.

In each case, there is an implied direct and immediate relationship between X and Y. To put it more formally, X and Y are directly linked in a formal logical train of action (first two examples) or thought (third example).

Now, how does this common sense meaning of the word “for” apply to the phrase “X is a gene for Y”? Let me illustrate the problem with a vignette

A jumbo jet contains about as many parts as there are genes in the human genome. If someone went into the fuselage and removed a 2-foot length of hydraulic cable connecting the cockpit to the wing flaps, the plane could not take off. Is this piece of equipment then a cable for flying?

Most of us would be uneasy answering yes to this question. Why? Because this example violates our conception of **causal proximity**. When we say X is for Y, we expect X to be, to a first approximation, directly and immediately related to Y. That is not the case for the cable and flying. There are many, many mechanical steps required to get from the function of that cable to a jumbo jet rising off the runway.

Another vignette:

Assume a Mendelian genetic disease due to a mutation in gene K. Gene K’s normal function is to produce an enzyme L that breaks down metabolite M in cells allowing M to be harmlessly secreted from the body. When K has a pathogenic mutation, the enzyme L that is produced no longer works. Therefore, levels of M rise, producing a well understood series of toxic effects, thereby producing the genetic disorder N.

This scenario suggests the following potentially simple causal chain: mutated gene K→dysfunctional enzyme L→excess metabolite M→disorder N. In this admittedly oversimplified story, a case could be made that gene K had sufficient causal proximity to disorder N to make plausible the claim that “K is a gene for N.” However, it might be argued that even here, the complexity of the paths from levels of M to disorder N may be far from “simple.”

Contrast this situation to the causal chain from a gene mutation to a complex psychiatric disorder such as schizophrenia. Although early efforts have been made to begin to trace such pathways (e.g., reference 56), we probably do not know enough to articulate all the specific causal steps that would be needed to go from DNA base-pair variation to, for example, the cognitive processes that predispose to delusion formation. What we can conclude with some confidence is that it will be very complex. Indeed, the causal link between that hydraulic cable and the jumbo jet flying will probably look very simple and short compared to the causal relationship between individual genes and the manifestations of schizophrenia. While the nature of the evidence reviewed here is largely inferential, it suggests that the pathways from most genes for psychiatric illness to their phenotypes would fail the causal proximity criterion implicit in the concept of “X is a gene for Y.”

**Appropriate Level of Explanation**

Scientific theories typically strive to explain phenomena at the most informative level. To provide an absurd example, no one would seek to understand the origin of hypertension at the level of quarks. In some ultimate way, quarks may be involved. But quarks are just the wrong level of inquiry for the problem.

To illustrate how this issue—the appropriateness of level of explanation—may apply to our evaluation of the concept of “a gene for...” consider these two “thought experiments”:

Defects in gene X produce such profound mental retardation that affected individuals never develop speech. Is X a gene for language?

A research group has localized a gene that controls development of perfect pitch (57). Assuming that individuals with perfect pitch tend to particularly appreciate the music of Mozart, should they declare that they have found a gene for liking Mozart?

For the first scenario, the answer to the query is clearly “No.” Although gene X is associated with an absence of language development, its phenotypic effects are best un-
derstood at the level of mental retardation, with muteness as a nonspecific consequence. X might be a “gene for” mental retardation but not language.

Although the second scenario is subtler, if the causal pathway is truly gene variant $\rightarrow$ pitch perception $\rightarrow$ liking Mozart, then it is better science to conclude that this is a gene that influences pitch perception, one of the many effects of which might be to alter the pleasure of listening to Mozart. It is better science because it is more parsimonious (this gene is likely to have other effects such as influencing the pleasure of listening to Haydn, Beethoven, and Brahms) and because it has greater explanatory power.

A final scenario:

Scientist A studied the behavioral correlates of a particular variant at gene X and concluded “This is a very interesting gene that increases the rates of sky diving, speeding, mountain climbing, bungee jumping, and unprotected casual sex.” Scientist B studied the same variant and concluded “This is a very interesting gene and effects levels of sensation-seeking.”

Who has done the better science? Since sensation seeking (and its close cousin novelty-seeking) are well studied traits (41), scientist B has provided results that are more parsimonious and potentially provide greater explanatory power. For example, only scientist B could predict that this gene ought to be related to other behaviors, like drug taking, that are known to be correlated with sensation-seeking.

As reviewed here, genes have been and will continue to be found that have statistical relationships with risk for psychiatric disorders. However, will the action of these genes be best explained at the level of the disorders themselves? While we cannot answer this question definitively, I would judge this to be unlikely. Far more plausible is that we will find genes whose mode of action can be best understood at the level of more basic biological processes (e.g., neuronal cell migrations during development) and/or mental functions (e.g., processing of threat stimuli).

Overview and Conclusion

The goal of this essay is to understand the historical origins of the key phrase “X is a gene for Y” and then to evaluate its appropriateness for psychiatric disorders. Our interest, of course, is not merely the phrase itself, but the conceptual framework that underlies this form of Genetalk. The use of the phrase “a gene for” implies (and in fact only makes sense in the context of) genes which—like preformationist anlagen—“code for” psychiatric illness in a simple, direct, and powerful way.

I argue that the concept of “a gene for...” can best be understood as deriving from preformationist developmental theory which, in turn, influenced the interpretation of the concept of a gene in the work of Mendel, in medical genetics, and most recently in human molecular genetics. Five criteria were proposed for evaluating whether the preformationist concept of “X is a gene for Y” is appropriate for psychiatric disorders. I then reviewed the available evidence, which was of variable quality, that addressed each of these criteria.

The strength of association between individual genes and psychiatric disorders is weak and often nonspecific. Genes do not appear to contain all the information needed for the development of psychiatric illness, since environmental factors have, for several disorders, been shown to have causal specificity. The action of genes on psychiatric disorders may frequently be contingent on environmental exposures, although much needs to be learned in this area. The causal chain from genes to psychiatric disorders is probably long and complex. The appropriate level of explanation for gene action is much more likely to be basic biological or mental processes that contribute to psychiatric disorders rather than the disorders themselves. Thus, with varying degrees of confidence, the genetic contribution to psychiatric disorders fails to meet any of the five criteria for the preformationist concept of “a gene for...”. The impact of individual genes on risk for psychiatric illness is small, often nonspecific, and embedded in causal pathways of stunning complexity.

On this basis, I suggest that we conclude that the phrase “X is a gene for Y,” and the preformationist concept of gene action that underlies it, are inappropriate for psychiatric disorders. The strong, clear, and direct causal relationship implied by the concept of “a gene for...” does not exist for psychiatric disorders. Although we may wish it to be true, we do not have and are not likely to ever discover “genes for” psychiatric illness.

References


The cerebellum means the “little brain” (when translated from Latin) and is nestled behind the “big brain” (i.e., the cerebrum) within the skull. The cerebellum traditionally was considered to be responsible primarily for the coordination of movement and motor learning. However, more recent anatomical and functional studies indicate that the cerebellum (especially in humans) plays a wider role in many cognitive functions, such as language, executive functions, and spatial cognition. In addition, a number of neurologic and psychiatric conditions have been associated with cerebellar dysfunction, including autism, attention deficit hyperactivity disorder, mood disorders, and schizophrenia. In humans, the cerebellum is a highly convoluted structure; in mice its gross structure is somewhat less complex (central part of Figure). However, in both humans and mice, the cellular organization is relatively simple. The cortex of the cerebellum is composed of three layers of neurons: the molecular layer, the Purkinje cell layer, and the granule cell layer. The cell bodies of the Purkinje neurons (in the Purkinje cell layer) extend long, elaborate, graceful dendrites into the molecular layer, as shown in the Golgi-filled Purkinje cell on the right. The major inputs into the cerebellum come from the inferior olive neurons and from neurons in the spinal cord and brain stem that send axons called mossy fibers into the cerebellar cortex. The mossy fibers synapse on the granule cells and the climbing fibers from olivary neurons wrap around the Purkinje cell dendrites as they climb into the molecular layer. The Purkinje cells are the only direct output neurons from the cerebellar cortex, sending inhibitory projections primarily to the deep cerebellar nuclei and a few extracerebellar regions, including the vestibular nuclei. Through indirect pathways, the cerebellum receives information from all sensory modalities, including auditory, visual, somatosensory, and proprioceptive systems as well as input from the neocortex. In turn, the cerebellum sends information indirectly throughout the brain. One important projection pathway leads to the thalamus and from there to the motor cortex. Functional neuroimaging studies in humans and physiological studies in rats and mice are revealing new insights into cerebellar function. One thing that is becoming clear is that the cerebellum plays a more complex role in how the brain functions than previously thought.

MICHAEL VOGEL, PH.D.
Baltimore, Md.

Address reprint requests to Dr. Tamminga, UT Southwestern Medical Center, Department of Psychiatry, 5323 Harry Hines Blvd., #NC5.914, Dallas, TX 75390-9070; Carol.Tamminga@UTSouthwestern.edu (e-mail).
In Sickness and Health

As a physician I’ve learned a great deal about the “doctor-patient” relationship, but only a fraction of what I’ve learned when I’m the one on the other side—the one in pain. That’s when I appreciate the exquisite sense of loss of control that patients experience. That is when I begin to understand the huge power that the healers, those “health care professionals,” hold over the ones in need of help. It’s pretty easy to get it when you’re lying on a stretcher waiting for the ride back from the X-ray department to the emergency room and the technicians are talking about their lunch break or their weekend plans and you’re looking at the ceiling tiles and wondering if the pain means something really serious—something that will change your entire life. It is also pretty clear when you’re the patient and the doctors come in for rounds and make a few statements and leave, and you wonder for the next 24 hours about the implications of those few words.

When I was sick this last time I had a new glimpse into that netherworld of the doctor-patient relationship. This time it wasn’t serious and I wasn’t in a hospital, just wracked with an intestinal virus that emptied my body of what seemed like all its fluids and strength. I was so depleted that my husband, who is also a physician, administered intravenous fluids to me at home over the course of an evening. Pretty unique actually, in an age when house calls by doctors have become so rare. The part about lying in my own bed next to a makeshift I.V. pole and using a glass jar from the kitchen for a “sharps container” as the fluid and medications were dripping into my arm was certainly a new and interesting esthetic experience, but the other part—the husband and wife as doctor and patient—was truly remarkable.

In our 20 years together we have seen each other sick, stressed, and vulnerable, and there have been occasions to care for one another. Physical and emotional intimacy are part of our relationship, with a reassuring ebb and flow of intensity. Nothing quite prepared me, however, for the psychic shock of watching my husband survey my arm for the purpose of finding the best vein to puncture. I was lying in bed—our bed—as he loomed over me. He held me with a tender and firm grip and his touch was gentle, but his mind was more distant—somewhere else—like a doctor. He completed his preparations, then penetrated the skin and the tissue below with a sharp needle. The act he has performed so many times on others but never on—or with—me. He secured the tubing and taped it in place, tidied up, and arranged my arm on a pillow, watching satisfied as the restorative fluid began to drip into my body.

“You were so nice to me,” I said in a slightly pathetic way, trying to convey the sense of appreciation of a helpless victim who is treated in a surprisingly kind manner by the often-cold doctor. “I treated you like I treat any other patient,” he replied. Strange answer. I guess I am glad that he is so nice to his patients, but maybe jealous too. On the other hand, I sensed that he was a bit distant with me—the way he needed to be to poke a hole in me with a sharp object—so I suppose he’s that way with them too. I could see and feel in a completely new way why patients fall in love with their doctors. The surge of gratitude and admiration that arises when someone in power focuses all his attention on you—on your health and well-being, on finding a way to minimize the pain of a helpful procedure. This is the unique physical and emotional intimacy of the doctor-patient relationship, which has the potential to be misconstrued as love, exploited by those in power, and cherished as an element of healing.

JEANNE STEINER, D.O.

Address correspondence and reprint requests to Dr. Steiner, Connecticut Mental Health Center, Yale School of Medicine, Rm. 148, 34 Park St., New Haven, CT 06508; jeanne.steiner@yale.edu (e-mail).
Milton H. Erickson, M.D., long considered the father of modern clinical hypnosis, is best appreciated today as a psychotherapy innovator (1). His “uncommon” therapy (2) was molded by his early career research into the nature of suggestion, hypnotic states, the mental mechanisms underlying psychodynamic processes, and the psychophysiological aspects of trance (3). Dr. Erickson had a pivotal realization that even as a hypnotherapist, he could be most effective when not using formal or directive hypnosis. The nondirective, naturalistic style he invented is called Ericksonian hypnosis, and his revolutionary psychotherapeutic approach is called Ericksonian psychotherapy. He conceptualized what he was doing as actively catalyzing some new possibility, not as passively awaiting change or as commanding, prescribing, or controlling the ultimate outcome. When I queried him about this in July 1977, he said, “Once you start a snowball rolling from the top of a mountain, who can tell what it will grow into and what path it will take?”

Dr. Erickson believed that treatment should be specifically constructed for each patient because each patient is unique. He advocated a “utilization” approach whereby a clinician utilizes whatever behavior, ideas, or attitudes patients exhibit. Dr. Erickson had a deeply humane view of patients, and he spent a great deal of time getting to know them while testing and working with their responses, especially their strengths. He would concoct an approach—not always using hypnosis—that would allow some slight change to occur, and this was often followed by a cascade of progressive changes. He appeared to have uncanny intuition yet attributed his clinical insight to both the acute development of his own perceptual skills and the innate receptive and synthetic capacities of the unconscious, which Dr. Erickson viewed as a reservoir of creative potential that can be a source of wisdom, not just of pathology.

Among the many therapeutic approaches he spawned were one-session therapy, brief therapy, strategic family therapy, systems-oriented therapy, solution-focused therapy, ordeals, and Ernst Rossi’s psychobiological therapy. He invented such therapeutic techniques as paradox, humor, reframing, confusion, surprise, binds and double binds, metaphors and storytelling, ambiguous function assignments, variable session length, and going with the resistance. Among the many hypnotic techniques he invented or developed are arm levitation, interspersal, confusion, illogic and irrelevance, indirection, many varieties of time distortion, age regression and progression, future projection, and trance induction by touch alone (1).

Dr. Erickson was a truly American character, born in Nevada, who came east in a covered wagon and grew up on a farm in Wisconsin. His struggles to overcome a nearly fatal bout of polio and, later, to live with postpolio syndrome shaped his therapeutic drive. Perhaps it is not surprising that this pioneering clinician who spent the last 30 years of his life practicing and teaching in Arizona and who used stories and metaphors frequently in his work with patients from near and far should himself be the subject of tales that sometimes portray him as larger than life.

Today, nearly 100 Erickson institutes exist in almost 30 countries. The American Society of Clinical Hypnosis, which Dr. Erickson cofounded along with its journal, is approaching its golden anniversary. So, too, is the decision by the American Medical Association to ratify hypnosis as a legitimate medical technique. Dr. Erickson was a courageous clinician whose ingenious insights and techniques can continue to enrich our therapeutic palette if only we are willing to listen to his voice.

References

GREGG E. GORTON, M.D.
Structural Brain Magnetic Resonance Imaging of Limbic and Thalamic Volumes in Pediatric Bipolar Disorder

Jean A. Frazier, M.D.
Sufen Chiu, M.D., Ph.D.
Janis L. Breeze, M.P.H.
Nikos Makris, M.D., Ph.D.
Nicholas Lange, Sc.D.
David N. Kennedy, Ph.D.
Martha R. Herbert, M.D., Ph.D.
Eileen K. Bent, B.A.
Vamsi K. Koneru, B.A.
Megan E. Dieterich, B.A.
Steven M. Hodge, M.A.
Scott L. Rauch, M.D.
P. Ellen Grant, M.D.
Bruce M. Cohen, M.D., Ph.D.
Larry J. Seidman, Ph.D.
Joseph Biederman, M.D.

Background: Youths with bipolar disorder are ideal for studying illness pathophysiology given their early presentation, lack of extended treatment, and high genetic loading. Adult bipolar disorder MRI studies have focused increasingly on limbic structures and the thalamus because of their role in mood and cognition. On the basis of adult studies, the authors hypothesized a priori that youths with bipolar disorder would have amygdalar, hippocampal, and thalamic volume abnormalities.

Method: Forty-three youths 6–16 years of age with DSM-IV bipolar disorder (23 male, 20 female) and 20 healthy comparison subjects (12 male, eight female) similar in age and sex underwent structured and clinical interviews, neurological examination, and cognitive testing. Differences in limbic and thalamic brain volumes, on the logarithmic scale, were tested using a two-way (diagnosis and sex) univariate analysis of variance, with total cerebral volume and age controlled.

Results: The subjects with bipolar disorder had smaller hippocampal volumes. Further analysis revealed that this effect was driven predominantly by the female bipolar disorder subjects. In addition, both male and female youths with bipolar disorder had significantly smaller cerebral volumes. No significant hemispheric effects were seen.

Conclusions: These findings support the hypothesis that the limbic system, in particular the hippocampus, may be involved in the pathophysiology of pediatric bipolar disorder. While this report may represent the largest MRI study of pediatric bipolar disorder to date, more work is needed to confirm these findings and to determine if they are unique to pediatric bipolar disorder.

Bipolar disorder is one of the most severe neuropsychiatric disorders at any age and is among the most disabling of psychiatric conditions that affect youths (1, 2). Although the chronology of underlying structural brain abnormalities in this population is unknown, such abnormalities may represent disruptions in typical brain growth resulting from an interplay of genetic and environmental factors. MRI studies are critical for advancing our knowledge of brain regions involved in the pathophysiology of pediatric bipolar disorder.

Youths with bipolar disorder are more severely ill and have higher genetic loading than adult-onset cases (3–5). Furthermore, children's brains are typically free from confounding factors known to affect brain structure and function (e.g., extensive treatment, substance use, history of electroconvulsive therapy). Finally, studying children and adolescents facilitates observation during a time in development when there are hormonal shifts known to have neuromodulatory effects on brain regions such as the temporal lobe (6, 7). By virtue of all of these factors, there is an increased likelihood of uncovering significant brain anatomic abnormalities in this early-onset group.

Although no discrete brain area has been consistently reported abnormal in the adult bipolar disorder MRI (structural) literature, some data lend support to several proposed neuroanatomic models of emotion regulation (8–10). One of these proposed models includes the following brain regions: prefrontal cortex, amygdala-hippocampus complex, hypothalamus, thalamus, insular cortex, ventral striatum, and interconnected structures (8). The amygdala, hippocampus, and thalamus are of particular interest in the study of bipolar disorder because of their functional roles in the brain. For example, the amygdala is integral in emotion-related aspects of behavior, memory, and learning; the thalamus processes sensory information and integrates activity among forebrain regions; and the hippocampus plays a role in learning and memory in providing contextual information (11). Several prior structural MRI studies of adults with bipolar disorder have reported abnormalities in the limbic structures and thalamus (12–18). Abnormalities in these structures might confer a propensity toward dysregulated mood states and vulnerability toward developing a mood disorder.

Prior structural MRI studies in pediatric bipolar disorder have indicated that there are anatomic abnormalities...
in a number of the structures implicated in the neural systems governing affective and cognitive processes. Botteron and colleagues demonstrated a loss of the normal asymmetry in the frontal lobe (19). A study conducted by Friedman and colleagues (20), which included adolescents with schizophrenia and bipolar disorder in a combined patient group, found that this patient group had reduced intracranial volumes and increased frontal and temporal sulcal sizes relative to healthy subjects. A third report that used the same subjects as in the Friedman et al. study found the patient group had reduced thalamic area relative to healthy subjects (21). A recent study consisting of a subgroup of 14 adolescents (age range=10–22 years). These youths had smaller hippocampal and amygdalar volumes than did 23 healthy adolescents (18). Finally, Del-Bello and colleagues (22) recently reported that adolescents with bipolar disorder (N=23) had smaller amygdala and enlarged putamen volumes compared with healthy subjects (N=20). In summary, prior MRI studies have suggested that youths with bipolar disorder have abnormalities in a number of the brain areas discussed by Soares and Mann (8) in their neuroanatomic model of emotion regulation: total cerebral volume, frontal lobe, hippocampus, amygdala, putamen, and thalamus. In order to assess anatomic findings in bipolar disorder further, we conducted a structural MRI study to evaluate brain volumes in early-onset bipolar disorder cases. We hypothesized that youths with bipolar disorder would have abnormalities in structures involved in the model of affect regulation discussed by Soares and Mann (8). In particular, the amygdala, hippocampus, and thalamus were chosen a priori on the basis of previous imaging studies of youths and adults with bipolar disorder.

Method

Subjects

The study was approved by institutional review boards at the Massachusetts General Hospital and McLean Hospital. Subjects were recruited through the McLean Hospital outpatient program and professional-patient advocacy groups. Inclusion criteria were DSM-IV diagnosis of bipolar disorder, age 6–16 years, and right-handedness. Male and female subjects of all ethnicities were recruited. Healthy subjects, all right-handed, were recruited through community newspaper advertisements and had no DSM-IV axis I diagnosis according to structured and clinical interviews and no family history of affective disorders or psychotic disorders in first-degree relatives. Exclusion criteria were major sensorimotor handicaps; full-scale IQ <70 or learning disabilities; history of claustrophobia, head trauma, loss of consciousness, autism, schizophrenia, anorexia or bulimia nervosa, electroconvulsive therapy, or alcohol or drug dependence/abuse (in the 2 months preceding the scan or a total history of 12 or more months); active medical or neurologic disease; metal fragments or implants; or current pregnancy or lactation.

Procedure

Seventy subjects (all outpatients) and their parents (or guardians) signed assent and informed consent forms. Three subjects were determined ineligible during interview, and one stopped the study because of lack of interest. Sixty-six scans were obtained; three scans were unreadable due to motion artifact. Therefore, data from 63 subjects scanned as part of an ongoing neuroimaging study are included in this report: 43 youths with DSM-IV bipolar disorder and 20 healthy subjects.

All of the youths underwent a diagnostic semistructured interview (Schedule for Affective Disorders and Schizophrenia for School-Age Children—Epidemiologic Version [K-SADS-E] [23]) and a clinical interview by board-certified child psychiatrists (J.A.F., S.C.). In addition, parents were administered an indirect K-SADS-E regarding their children by trained raters. These B.A.-level raters received 4 months of training on the administration of the K-SADS-E under the supervision of senior raters and the senior investigator (J.B.). All raters had established a high degree of interrater reliability: from 175 interviews, the mean kappa was 0.90, and all disorders achieved kappa coefficients >0.82. Final DSM-IV diagnoses were established by the consensus diagnosis of clinical and structured interviews.

Each youth received a physical and neurological examination that included Tanner staging (a 1–V scale of pubertal development) (24) and cognitive testing. The age at onset of each illness was determined by parental report of symptoms on the structured interview. Age at illness onset was defined as the time when the youth met full diagnostic criteria (e.g., age at onset of bipolar illness was the age at which the youth first met full diagnostic criteria for mania). Children and adolescents were given several subtests of the Wechsler Intelligence Scale for Children, 3rd ed. (WISC-III) (25) which permitted the estimation of verbal IQ. Handedness was assessed using the Edinburgh Handedness Questionnaire (26).

Measures of current psychopathology were obtained using the Young Mania Rating Scale (27) and Global Assessment of Functioning Scale (GAF) (DSM-IV, p. 32).

Antipsychotic doses (converted to chlorpromazine equivalents) (28, 29), as well as number and type (antipsychotic, antidepressant, stimulant, anticonvulsant, lithium) of psychoactive medications at the time of scan were used as clinical variables.

MRI Protocol

Structural imaging was performed at the McLean Hospital Brain Imaging Center on a 1.5-T Signa scanner (GE Medical Systems, Milwaukee). Acquisitions included a conventional T1-weighted sagittal scout series (20 slices), a proton density/T2-weighted interleaved double-echo axial series (120 slices, slice thickness=3 mm, field of view=24 cm², TR=3 seconds, TE=30/80 msec, acquisition matrix=256×192, number of excitations=0.5), and a three-dimensional inversion recovery-prepped spoiled gradient recalled echo coronal series, which was used for structural analysis (124 slices, prep=300 msec, TE=1 minute, flip angle=25°, field of view=24 cm², slice thickness=1.5 mm, acquisition matrix=256×192, number of excitations=2). All scans were reviewed by a clinical neuroradiologist to rule out gross pathology.

Image Analysis

Structural scans were transferred to the NMR Center for Morphometric Analysis-Charlestown Massachusetts General Hospital and coded and catalogued for blind analysis. Imaging analysis was done on Sun Microsystems, Inc. (Mountainview, Calif.) workstations using Cardviews software (30). The datasets were positionally normalized to overcome variations in head position by imposing a standard orientation on each scan using the midpoints of the decussations of the anterior and posterior commissure lines and the midsagittal plane at the level of the posterior commissure as points of reference for rotation and translation. The images were not rescaled to Talairach spatial dimensions in order...
to preserve individual and interhemispheric differences in the morphometry of structures (31).

The entire image sets were then segmented into gray, white, and CSF tissue classes by three image analysts, under the supervision of one of the authors (N.M.), all of whom had strong backgrounds in neuroanatomy and extensive training in morphometric analysis and who were blind to subject-identifying information. The segmentation method uses a semiautomated intensity contour algorithm for external border definition and signal intensity histogram segmentation method uses a semiautomated intensity contour algorithm for external border definition and signal intensity histogram distributions for delineation of gray-white borders (Figure 1). This technique allows for border definition as the midpoint between the peaks of the bimodal distribution for any given structure and its surrounding tissue (32).

Total cerebral volume. Segmentation of the regions of interest was performed following the anatomic definitions of Filipek and colleagues (33) for the total cerebrum. Total cerebral volume was defined as all gray and white matter in the cerebrum and did not include CSF; cerebellum, or brain stem. The cerebrum was measured across all 124 coronal slices in which it appeared.

Thalamus. Segmentation of the thalamus was performed following the anatomic definitions of Seidman and colleagues (34, 35) by tracing the trajectory of the hypothalamic fissure in the sagittal plane to separate the thalamus proper from the ventral diencephalon. The medial boundary of the structure was the third ventricle, and the lateral boundary was the internal capsule. The superior border was the body of the lateral ventricle, and the inferior border was the hypothalamic fissure.

Hippocampus and amygdala. The method of Filipek et al. (33) defines the amygdala and hippocampus as a continuous gray matter structure in the primary segmentation. These two structures are then separated from each other according to the procedure described by Seidman and colleagues (36) in which the hippocampus is separated from the amygdala at the rostral-coronal plane, where the hippocampus first appears. The segmentation of the amygdala was performed manually in its entirety, comprising approximately 11 subsequent coronal sections.

The coexistence of the amygdala and hippocampus in several coronal sections can make the precise identification of the ventral amygdalar border difficult (37). Therefore, we used the cross-referencing capability of the program Cardview (38) to draw outlines delimiting the amygdala in axial and sagittal views (Figure 2 B, C); this preliminary procedural step allows a reliable separation of the amygdala from surrounding gray structures, such as the ventral part of the lentiform nucleus, the medial temporal cortex, and the hippocampus, thus eliminating the need to apply conventions to define the anterior amygdalar boundary (12, 22) or the amygdala-hippocampal junction (12). The anterior portion of the amygdala was segmented as it appears beneath the medial temporal cortex (slice K in Figure 2). At this region, the medial temporal cortex and the amygdala can give the impression of being only a thickening of the medial temporal cortex with no amygdala present, as has been reported previously (12, 22). Therefore, the definition of these borders was particularly aided by the tracing of cross-referenced outlines in the axial and sagittal planes (Figure 2 B, C). The choroidal fissure was used as the superior border of the amygdala along with the gray-white matter contrast between the amygdala and surrounding white matter. The gray-white matter contrast between the amygdala and its surrounding temporal white matter (consisting of the centrally located temporal white matter stem), as well as the gray-CSF contrast between the amygdala and the temporal horn of the lateral ventricle, was considered the lateral border of the amygdala. Finally, the inferior border consisted of the gray-white matter contrast between the amygdala and its surrounding temporal white matter anteriorly and by the alveus (of the hippocampus) and the temporal horn of the lateral ventricle posteriorly.

The volumes for each structure were derived by multiplying the number of voxels assigned to each structure on each slice by the voxel volume (the product of slice thickness and the square of in-plane resolution), followed by summing across all slices in which the structure appeared; the volumes are reported in cm³ (31). For the reliability study, 10 scans from our data set were selected at random and blindly segmented by two raters. Five of the scans were also remeasured in a random order by one of the raters to estimate the intraclass correlation coefficient.

The standard interrater intraclass correlation coefficient for the total cerebrum was 0.93, and the intra- and interrater correlation coefficients, respectively, for the regions of interest were 0.88 and 0.84 for the amygdala, 0.95 and 0.96 for the thalamus, and 0.93 and 0.94 for the hippocampus.

Data Analyses

S-Plus 6.0 (Insightful Corp., Seattle) was used for statistical analysis. All statistical tests were two-sided with alpha set at 0.05. Differences in demographic and clinical variables were measured using t tests for continuous variables and chi-square tests for categorical variables. In addition, Pearson’s correlations were calculated for clinical variables of the bipolar disorder group and those structures that differed significantly between bipolar disorder youths and healthy subjects. The clinical variables included Young Mania Rating Scale and GAF scores, number of psychoactive medications, chlorpromazine equivalents, and verbal IQ.

Noting that brain structure sizes possess variances that increase with their means, we analyzed our volumetric data on the natural logarithmic scale as a step toward uncoupling this non-Gaussian relationship. We conducted an exploratory multivariate analysis of variance (MANOVA) on the three-dimensional ensemble of log total hippocampal, amygdalar, and thalamic volumes with the effects of log total cerebral volume, age, sex, diagnosis, and the sex-by-diagnosis interaction controlled. In addition, an exploratory MANOVA analysis was performed to assess the effects of age group, mood state, medication type, and presence of ADHD or psychosis on the log hippocampal and log total cerebral volumes.

We then proceeded to analyze the effects of these covariates on log brain structure sizes by three separate univariate linear regres-
FIGURE 2. Segmentation of the Amygdala, Hippocampus, and Thalamus

The segmentation method of the amygdala, hippocampus, and thalamus is shown in T₁-weighted MR images. In particular, the segmentation method used for the amygdala (developed by N.M.) is shown in detail in images J–P. Image A shows a sagittal slice passing through the hippocampus and amygdala. Yellow vertical lines show the locations of representative coronal sections J–Q. The blue box in A indicates the region containing amygdala (K–N), amygdala-hippocampal transition, anterior hippocampus (L), and the temporal-polar region rostral to the amygdala (J). B shows an axial section (the position of which is indicated by arrow in A) used for the tracing of an outline to separate the most anterior portion of the amygdala from the surrounding medial temporal cortex (red line). C shows an enlarged version of the blue box in A to emphasize the amygdala-hippocampal transition. In both B and C, the light green line represents the border between amygdala and hippocampus. In C, the red line shows the superior, anterior, and inferior limits of the amygdala. These outlines are used as guidelines to complete segmentation on the coronal plane (D–J). Similarly, the thalamus is segmented based on the intensity contrast between this structure and its surrounding white matter; a border delimits its inferior border at the hypothalamic fissure. Structural details from coronal slices J–O are enlarged to show the method of amygdala segmentation. J shows a rostral slice in the temporal lobe where the amygdala is not present. More caudally, the anterior part of the amygdala appears within the rostromedial temporal area beneath the cortex as shown in K. Further posteriorly, the middle portion of the amygdala occupies a position superior and lateral to the anterior parahippocampal gyrus as shown in L. Progressing in the rostrocaudal dimension, the hippocampus appears as shown in M and N, in which the amygdala is also present. In a more posterior location (O), there is only hippocampus as amygdala is no longer present. In P, the posterior most segment of the hippocampus is shown flanked laterally by the atrium of the lateral ventricle. Finally, Q is the coronal posterior to the hippocampus. To emphasize the amygdala and hippocampal outlines, the segmentation outlines of the other brain structures were omitted from images J–Q.
Results

Data from 63 subjects are included: 43 youths with DSM-IV bipolar disorder (mean age=11.3 years, SD=2.7; current episode: mixed=52.3%, manic=15.9%, depressed=11.4%, euthymic=20.5%) and 20 healthy comparison subjects (mean age=11.0 years, SD=2.6). Characteristics of the groups are summarized in Table 1. There were no significant height, head circumference, or Tanner stage differences between groups; there was a significant difference in weight (Table 1).

Although the bipolar disorder youths had verbal IQ scores in the normal range, they were significantly lower than those of comparison subjects (mean=100.7 [SD=14.3] versus 116.9 [SD=12.7]; t=4.3, df=61, p<0.001). The bipolar disorder group also scored lower on the GAF (mean=49.5 [SD=6.0] versus 68.8 [SD=1.9]; t=19.0, df=61, p<0.001). The bipolar disorder subjects had a number of comorbid conditions (mean=7.2, SD=3.2). The most common comorbid conditions were oppositional defiant disorder (67% [N=29]) and ADHD (51% [N=22]). Most bipolar disorder youths (86% [N=37]) had experienced at least one episode of major depression (mean age at onset=6.8 years, SD=3.6), and 40% (N=17) had a history of psychosis. One youth had a history of alcohol abuse, which occurred more than 3 months before enrollment. Medications used at the time of MRI included lithium (26% [N=11]), anticonvulsants (42% [N=18]), antidepressants (30% [N=13]), stimulants (21% [N=9]), atypical antipsychotics (76% [N=33]), and other (19% [N=8]), which included anticholinergics and beta-adrenergics. Table 2 presents details of the clinical and treatment characteristics of the bipolar disorder patients.

Clinical neuroradiological interpretations of the scans showed normal variants in three healthy subjects (slightly prominent lateral ventricles [N=1], large cisterna magna [N=1], and pineal cyst [N=1]) and two bipolar disorder subjects (mildly prominent lateral ventricles [N=1] and large cisterna magna [N=1]). One healthy subject had findings that were atypical but of unclear clinical significance (a tiny focus of hypointensity in the subcortical white matter of the superior left frontal lobe). Nine bipolar disorder subjects had findings that were atypical and again of unclear clinical significance (prominent ventricles [N=3], right-greater-than-left asymmetry of the temporal horn [N=2], multiple white matter hyperintensities in the bilateral parietal area [N=1] and in the left hemisphere [N=1], nonspecific punctate T2 foci in the left parietal region [N=1], and bilateral widened perivascular spaces noted in the inferolateral portion of the basal ganglia [N=1]).

Volumetric Measurements

Table 3 contains the limbic and thalamic volume data. The amygdala volumes were measured across a mean of 12.1 slices (SD=1.6, range=9–15) on the right side, and 11.5 slices (SD=1.4, range=8–15) on the left. The hippocampal volumes were measured across a mean of 24.4 slices (SD=1.7, range=21–28) on the right side, and 24.7 slices (SD=2.0, range=20–29) on the left. The thalamic volumes were measured across a mean of 22.7 slices (SD=1.3, range=18–25) on the right side, and 22.9 slices (SD=1.1, range=21–25) on the left.

No significant correlations were seen between any clinical variables and the hippocampus or total cerebral volume for the bipolar disorder group.

When the bipolar disorder group was assessed using the exploratory MANOVA analysis, we found no significant effects of age group, mood state, medication type, or presence of ADHD or psychosis on volume in the hippocampus or cerebrum.

The initial exploratory MANOVA of the three-dimensional brain structure ensemble yielded potentially significant effects for sex (F=2.76, df=3, 55, p=0.05) and the sex-by-diagnosis interaction (F=2.17, df=3, 55, p=0.10).

Hippocampus. The linear regression model for log total hippocampal volume contained the same effects as in the preliminary MANOVA. We found a possible sex-by-diagnosis interaction in raw hippocampal volumes by sex and diagnosis (Table 3). Although not significant, we retained age in the model, since this covariate was fixed by our study design; main effects and the interaction of sex and diagnosis remained significant when age was dropped. We tried controlling for height and weight but found that neither of their effects was significant; the sex-by-diagnosis interaction remained significant. We found significant effects of log cerebral volume, sex, diagnosis, and the sex-
by-diagnosis interaction (Table 4). We pursued the source of interaction by fitting separate regression models of the same form to the boys and the girls. This step confirmed our pooled variance assumption when fitting the single model with a sex effect (39). We conclude that the sex-by-diagnosis interaction is being driven by the smaller hippocampus of the girls after log total cerebral volume and age are controlled.

**Amygdala and thalamus.** The same logarithmic transformations and model selection procedure was applied to the thalamic and amygdalar volumes separately. No significant effects of sex, diagnosis, or their interaction were found.

**Discussion**

The youths in our study with bipolar disorder had significantly smaller total hippocampal and total cerebral volumes. These findings contribute to the existing literature on bipolar disorder youths, which has shown loss of normal frontal lobe asymmetry (19), reduced intracranial volume (20), increased frontal and temporal sulcal size (20), reduced thalamic area (21), larger putamen volume (12, 22), and smaller amygdala (18, 22) and hippocampal (18) volumes relative to healthy subjects.

These findings in children and adolescents differ somewhat from those in adults, but comparison of MRI results across studies is difficult because of variations in methodology. For example, until recently, it was difficult to even reliably measure structures such as the hippocampus—and particularly the amygdala—on MRI scans (12, 22). It is of note that the quality of the T1-weighted images used in the present study, as well as the outlines traced on the cross-referenced axial and sagittal slices, allow reliable visualization and segmentation of the amygdala, eliminating the need to apply conventions for defining either the anterior amygdalar boundary (12, 22) or the amygdala-hippocampal junction (12). Despite the differences in the methodologies used across studies, it is worth highlighting the anatomic variations that have been reported in adult structural MRI studies relative to the findings reported here and in other pediatric studies of bipolar disorder patients.

Of the structures included in our a priori hypothesis, the thalamus and amygdala have varied findings in adult and child and adolescent studies. Several adult bipolar disorder studies have reported that both structures are in-
NEUROIMAGING FINDINGS IN PEDIATRIC BIPOLAR DISORDER

Our finding of decreased hippocampal volumes is the first report of this finding in children and adolescents with bipolar disorder and further extends the findings of Blumberg and colleagues in which bilateral hippocampal volume reductions were seen in adolescents with bipolar disorder but not in adults with the illness (18). Adult bipolar disorder studies of hippocampal volumes have found either decreased or normal volumes (12, 16, 18, 41). It is possible that the adult studies in which reduced hippocampal volumes were reported may have included at least some adults who had childhood-onset illness. However, this information was not included in those studies. If age at onset, particularly childhood-onset, is related to reduced hippocampal volumes, this could partially explain the mixed results seen in adult studies. Our finding of reduced hippocampal volume, as well as another group’s similar finding among adolescents with the illness (18), may represent a finding unique to early-onset bipolar disorder. This finding in our youths is of interest in light of neuropathologic studies of the hippocampus in bipolar disorder, which have suggested there may be abnormal neurodevelopment and remodeling of synapses in that structure (42–48).

A study of adults that compared monozygotic twins discordant for bipolar disorder found that the right hippocampus was smaller in the sick twin compared with the well twin (17), suggesting that this finding might be a structural correlate for the presence of disease. Given that reduced hippocampal volume was the only significant limbic finding in our group of bipolar disorder youths, and that at least two studies in adults (16, 17) had a similar finding in the right hippocampus, this may also be a finding reflective of disease.

Female bipolar disorder subjects showed a more pronounced decrease in hippocampal volume than did male bipolar youths relative to their comparison subjects. This sex-by-diagnosis interaction could reflect a variety of factors. For example, girls may need to have more significant structural abnormalities in order to reach threshold for disease expression. Alternatively, since the girls with bipolar disorder in our study had significantly higher Young Mania Rating Scale scores than the boys, the smaller hippocampal volumes in girls might be reflective of a greater degree of psychopathology. However, both male and female bipolar disorder youths had reduced hippocampal volumes, and it may be that there is some abnormality in the hippocampus, reflective of disease, that expresses itself more robustly in girls than in boys (49).

In our study we found that total cerebral volume was smaller in bipolar disorder youths (mean=5.4%, SD=6.4%); smaller total cerebral volume has also been reported in four prior studies of bipolar disorder youths (19–21, 50). Most adult bipolar disorder studies, including the recent study published by Blumberg and colleagues (18), have not found differences in total cerebral volume. Therefore, the difference seen in total cerebral volume between early- and adult-onset bipolar disorder cases relative to healthy groups suggests that affected youths may have distinct neurodevelopmental trajectories, possibly from having brains that have developed differently or from early apopotic pruning of neuronal circuits. A smaller total cerebral volume in early-onset bipolar disorder illness may be reflective of a neurodevelopmental phenomenon. It should be noted that there is evidence of nonuniform scaling

### TABLE 4. Total Cerebral Volume, Age, Sex, Diagnosis, and Sex-by-Diagnosis Interaction Effects on Limbic and Thalamic Volumes in Pediatric Patients With Bipolar Disorder and Healthy Comparison Subjects

<table>
<thead>
<tr>
<th>Brain Area and Covariate</th>
<th>Estimated Effect</th>
<th>SE</th>
<th>t (df=61)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amygdala</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log total cerebral volume</td>
<td>0.872</td>
<td>0.272</td>
<td>3.209</td>
<td>0.002</td>
</tr>
<tr>
<td>Age</td>
<td>0.002</td>
<td>0.007</td>
<td>0.229</td>
<td>0.82</td>
</tr>
<tr>
<td>Sex</td>
<td>0.019</td>
<td>0.064</td>
<td>0.299</td>
<td>0.77</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>-0.013</td>
<td>0.062</td>
<td>-0.206</td>
<td>0.84</td>
</tr>
<tr>
<td>Sex-by-diagnosis</td>
<td>-0.003</td>
<td>0.076</td>
<td>-0.040</td>
<td>0.97</td>
</tr>
<tr>
<td>Hippocampus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log total cerebral volume</td>
<td>1.068</td>
<td>0.150</td>
<td>7.134</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age</td>
<td>-0.001</td>
<td>0.004</td>
<td>-0.212</td>
<td>0.84</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.090</td>
<td>0.035</td>
<td>-2.558</td>
<td>0.02</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>-0.086</td>
<td>0.034</td>
<td>-2.529</td>
<td>0.02</td>
</tr>
<tr>
<td>Sex-by-diagnosis</td>
<td>0.100</td>
<td>0.042</td>
<td>2.385</td>
<td>0.02</td>
</tr>
<tr>
<td>Thalamus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log total cerebral volume</td>
<td>0.593</td>
<td>0.097</td>
<td>6.143</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age</td>
<td>0.005</td>
<td>0.003</td>
<td>1.832</td>
<td>0.08</td>
</tr>
<tr>
<td>Sex</td>
<td>0.022</td>
<td>0.023</td>
<td>0.957</td>
<td>0.35</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>-0.024</td>
<td>0.022</td>
<td>-1.098</td>
<td>0.28</td>
</tr>
<tr>
<td>Sex-by-diagnosis</td>
<td>0.022</td>
<td>0.027</td>
<td>0.807</td>
<td>0.43</td>
</tr>
</tbody>
</table>
across the brain regions studied herein relative to global changes in total cerebral volume (51).

Although abnormal anatomy does not necessarily confer abnormal function, some structural differences may increase the risk of developing dysfunction. Such differences might also be markers of preexisting susceptibility or vulnerability and could aid in identifying different patient phenotypes (52). Bipolar disorder youths often have behavioral and developmental difficulties early in life (2, 52–54), and their symptoms of bipolar disorder typically begin between the ages of 5–11 years (1, 3, 55); in our study the mean age at onset was 7.0 years (SD=3.8). Many of the progressive and regressive events in the brain, particularly in the temporal lobe, occur during this age range, and alterations in normal brain development during this time may result in the symptoms of bipolar disorder.

Adult genetic and neuroimaging studies in both schizophrenia and bipolar disorder lend support for a multifactorial etiology; one possibility is of a “two-hit” hypothesis, a genetic predisposition in combination with environmental influences, resulting in the disorder (35, 36, 56, 57). For example, our finding of smaller hippocampal volumes may reflect developmental or genetic influences or environmental insults (e.g., hypoxia) that have led to reductions in what were previously normally developing volumes. This hypothesis suggests that two hits (genetic and environmental), occurring either independently or in interaction, may result in reduction of hippocampal volumes as well as in the expression of bipolar disorder.

This report describes to our knowledge the largest MRI study of pediatric bipolar disorder to date. However, the number of subjects was still relatively small and represents only a cross-sectional look at youths with bipolar disorder and comparison subjects. Our findings should be considered in light of other limitations, such as the smaller number of comparison subjects relative to the number of youths with bipolar disorder, the lack of rating scales for depressive symptoms, and the difficulty in reliably determining the age at onset of a child or adolescent’s mood instability based on parental recall. Although we did examine the effects of several clinical variables, including age and type of medication as well as chlorpromazine equivalents, and found no significant effects, the power in our study may have been insufficient to fully assess the possible effects of these parameters.

Our findings support the hypothesis that the hippocampus may be involved in the underlying pathophysiology of pediatric bipolar disorder and may represent a unique characteristic of early-onset presentation of the disorder due to a genetic diathesis or derangements in growth processes around the time of illness onset. However, in order to obtain a better sense of the trajectory of abnormalities in this structure in bipolar disorder populations, longitudinal studies are needed. Future studies need large enough sample sizes of both sexes and a sex-matched healthy comparison group to sort out diagnostic, age at onset, developmental, and sexual influences on these structures over time (6, 38, 58). A larger group of subjects is currently being accrued in order to further assess these findings with greater power and to assess the sexual dimorphism of other brain structures.

Received June 24, 2003; revisions received Dec. 29, 2003, and July 9, 2004; accepted July 19, 2004. From the Department of Psychiatry and the Neuroscience Program, Harvard Medical School, Boston; the McLean Hospital Child Psychiatry Outpatient Clinic, Brain Imaging Center, Statistical Neuroimaging Laboratory, and Developmental Neuroscience Laboratory; the Departments of Psychiatry, Pediatric Neurology, Radiology, Neurology, and Pediatrics, the Pediatric Psychopharmacology Unit, and the Center for Morphometric Analysis, Massachusetts General Hospital, Boston; the Department of Psychiatry and Behavioral Sciences, University of California, Davis; the Department of Biostatistics, Harvard School of Public Health, Boston; and the Department of Psychiatry, Massachusetts General Hospital Child Psychiatry Outpatient Clinic, Boston. Address correspondence and reprint requests to Dr. Frazier, Child and Adolescent Neuropsychiatric Research Program, Cambridge Health Alliance/Mystic Center, 1493 Cambridge St., Cambridge, MA 02139; frazier@challiance.org (e-mail). Supported by NIMH research grants to Dr. Frazier (K08 MH-01573-01) and Dr. Lange (NS-37483).

The authors thank Mary Ahn, M.D., Sandra DeJong, M.D., and Jay Giedd, M.D., for editorial comments and Jill Garroway, Rebecca Melrose, Nathan Stein, Mari Sohima, and Shuna Klviness for their assistance with the study.

References


Am J Psychiatry 162:7, July 2005


21. Dasari M, Friedman L, Jesberger J, Stuve TA, Findling RL, Swales TP; Schulz SC: A magnetic resonance imaging study of thalamic area in adolescent patients with either schizophrenia or bipolar disorder as compared to healthy controls. Psychiatry Res 1999; 91:155–162


23. Orvaschel H, Puig-Antich J: Schedule for Affective Disorders and Schizophrenia for School-Age Children—Epidemiologic Version (K-SADS-E), 4th revision. Fort Lauderdale, Fla, Nova University, Center for Psychological Studies, 1987


44. Fatemi SH, Erle JA, McMenomy T: Reduction in Reelin immunoreactivity in hippocampus of subjects with schizophrenia, bipolar disorder and major depression. Mol Psychiatry 2000; 5:654–663


Familial Variation in Episode Frequency in Bipolar Affective Disorder

Maria E. Fisfalen, M.D.
Thomas G. Schulze, M.D.
J. Raymond DePaulo, Jr., M.D.
Leslie J. DeGroot, M.D.
Judith A. Badner, M.D.
Francis J. McMahon, M.D.

Objective: Bipolar affective disorder is a familial illness characterized by recurrent episodes of mania and depression, but little is known about the familial nature of episode recurrence or its associated clinical features. The authors analyzed the recurrence frequency of affective episodes (episode frequency), along with associated clinical and demographic variables, in families with at least three members with a major affective disorder.

Method: Members of 86 families ascertained through probands with bipolar affective disorder who had two or more first-degree relatives with a major affective disorder were interviewed by psychiatrists and assigned an all-sources diagnosis. Data for 407 subjects with a major affective disorder were analyzed. Episode frequency was estimated as the number of episodes of major depression, mania, and hypomania per year of illness.

Results: Episode frequency was smoothly distributed over the range of 0.02–20.2 episodes/year. Episode frequency was significantly correlated among relatives (r=0.56, p<0.004). Earlier age at onset, bipolar II disorder, hallucinations or delusions, alcoholism, and suicidal behavior were all more prevalent in the highest than in the lowest quartiles of episode frequency. Female gender and recurrent major depression were more prevalent in the lowest quartile. Panic disorder, substance abuse, and thyroid disease were all unrelated to episode frequency. Subjects with DSM-IV rapid cycling did not differ from other affected subjects for most of the variables tested.

Conclusions: Episode frequency is a highly familial trait in bipolar affective disorder, associated with several indicators of severity, and may be useful in defining clinical subtypes of bipolar affective disorder with greater genetic liability. DSM-IV rapid cycling was not supported by these data as the best predictor of familiality or severity.

All patients with bipolar affective disorder experience recurrent episodes of major depression, mania, hypomania, or mixed states, but the frequency with which episodes recur can range from one in several years to many per day (1). As many as 20% of people with bipolar affective disorder experience rapid cycling, defined by DSM-IV as four or more affective episodes in a year (1, 2). Rapid cycling is more common in women and in people with bipolar II disorder (1, 3–5). Hypothyroidism, steroid hormones, and the use of antidepressants have been associated with rapid cycling, but these findings are controversial (6–10). Rapid cycling is just one extreme of the spectrum of episode frequency. Few studies have examined the full range of episode frequency in bipolar affective disorder.

Similarly, although bipolar affective disorder is a familial illness, the nature of episode frequency as a familial trait has not been investigated extensively (1, 11). Several studies have addressed morbid risk for bipolar disorder among the relatives of rapid-cycling probands (4, 5, 12, 13), and some studies have addressed the tendency of rapid cycling to run in families (14, 15), but we are aware of no previous studies that directly measure the familiality of episode frequency. This issue is important, because the absence of familiality could suggest that episode frequency is under the primary control of nongenetic factors. We report here an analysis of episode frequency in families ascertained for a genetic linkage study of bipolar affective disorder. We also examined the relationship between episode frequency and age at onset, suicidal behavior, psychosis, panic disorder, alcoholism, substance abuse, and thyroid disease in these families. We found that episode frequency is a highly familial trait, associated with several indicators of severity, that may help define clinical subtypes of bipolar affective disorder with the greatest genetic liability.

Method

Subjects

A total of 625 subjects in 86 families ascertained for a genetic linkage study (16) were eligible for the current analysis. All study volunteers gave written informed consent after the procedures had been fully explained.

Families were ascertained through a proband with a reported history of bipolar I disorder and at least two first-degree relatives (at least two siblings or at least one sibling and only one parent) with a major affective disorder.

http://ajp.psychiatryonline.org
TABLE 1. Characteristics of Probands With Bipolar I, II, and Schizoaffective Bipolar Disorder (N=86) and Their Relatives With a Major Affective Disorder (N=321) in a Study of Familial Variation in Episode Frequency, by Diagnostic Group

<table>
<thead>
<tr>
<th>Diagnostic group</th>
<th>Gender</th>
<th>Age at Onset (years)</th>
<th>Age at Interview (years)</th>
<th>Duration of Illness (years)</th>
<th>Number of Episodes of Hypomania</th>
<th>Number of Episodes of Mania</th>
<th>Number of Episodes of Depression</th>
<th>Number of Episodes per Yeara</th>
</tr>
</thead>
<tbody>
<tr>
<td>All diagnoses (N=407)</td>
<td>146</td>
<td>261</td>
<td>22.1</td>
<td>18.5</td>
<td>10.4</td>
<td>40.9</td>
<td>10.4</td>
<td>40.9</td>
</tr>
<tr>
<td>Bipolar I disorder (N=144)</td>
<td>59</td>
<td>85</td>
<td>21.1</td>
<td>20.6</td>
<td>10.9</td>
<td>27.8</td>
<td>11.7</td>
<td>10.1</td>
</tr>
<tr>
<td>Bipolar II disorder (N=151)</td>
<td>57</td>
<td>94</td>
<td>20.4</td>
<td>23.7</td>
<td>12.6</td>
<td>47.9</td>
<td>41.1</td>
<td>17.3</td>
</tr>
<tr>
<td>Recurrent unipolar depression (N=102)</td>
<td>23</td>
<td>79</td>
<td>25.9</td>
<td>17.5</td>
<td>11.8</td>
<td>—</td>
<td>—</td>
<td>5.7</td>
</tr>
<tr>
<td>Schizoaffective disorder, bipolar type (N=10)</td>
<td>7</td>
<td>3</td>
<td>22.4</td>
<td>40.9</td>
<td>12.9</td>
<td>—</td>
<td>—</td>
<td>5.6</td>
</tr>
</tbody>
</table>

a Significant overall difference among diagnostic groups (p<0.0001, analysis of variance). In paired comparisons, significant differences between bipolar II disorder and bipolar I disorder groups (t=7.8, df=192, p<0.0001), bipolar II disorder and recurrent unipolar depression groups (t=10.5, df=165, p<0.0001), bipolar II disorder and schizoaffective disorder, bipolar type, groups (t=6.4, df=30, p<0.0001), and bipolar I disorder and recurrent unipolar depression groups (t=5.5, df=222, p<0.0001).

b Significant difference between the proportion of male and female subjects (χ²=10.5, df=1, p=0.001).

FIGURE 1. Number of Depressive and Manic Episodes per Year Since Onset of Major Affective Disorder in Probands With Bipolar I, II, and Schizoaffective Bipolar Disorder (N=86) and Their Relatives With Major Affective Disorder (N=321)²

Among affected subjects, the average level of education was 14.3 years (SD=3.0), and 77.3% were employed. Twenty-six percent were married, 55.8% were separated or divorced, and the rest were widowed or never married.

Clinical Assessment

All subjects were interviewed by a psychiatrist who used the Schedule for Affective Disorders and Schizophrenia—Lifetime Version (SADS-L) (17). Two additional psychiatrists reviewed the interview notes, family informant data, and medical records before assigning best-estimate diagnoses under Research Diagnostic Criteria (18). The diagnosis of bipolar II disorder required recurrent major depression as well as hypomania. All diagnoses were found to be highly reliable, on the basis of assessments of co-rated and test-retest interviews, as well as agreement between the psychiatrists who provided the best-estimate diagnoses (19). Final diagnoses among the probands were as follows: bipolar I disorder, N=71; bipolar II disorder, N=12; and schizoaffective disorder, bipolar type, N=3. In the total study group of probands and their family members, 23.0% (N=144) had bipolar I disorder, 24.2% (N=151) had bipolar II disorder, 16.2% (N=102) had recurrent unipolar depression, 1.6% (N=10) had schizoaffective disorder, bipolar type, 7.0% (N=44) had an uncertain diagnosis, and 27.8% (N=174) were deemed unaffected. Of the 407 affected subjects, 64% were female.

Retrospective lifetime self-report data for several clinical variables were collected with the SADS-L. This study focused on age at onset and number of episodes of mania, hypomania, and major depression, which were derived directly from the corresponding items in the SADS-L instrument. Age at onset was defined as age at first mania or major depression, whichever was earlier (20). The number of episodes was estimated by the interviewer, who worked with each subject to define in detail the most severe episode of mania, hypomania, and major depression, then asked, “How many episodes like this have you had in your lifetime?” Subjects were asked to count periods of illness separated by at least 2 months of recovery as separate episodes. Continuous periods of illness involving a switch in polarity were counted as two episodes, provided that each pole appeared to fulfill the diagnostic criteria, but these switches were uncommon.

Data were also extracted for six clinical variables that have been associated with bipolar affective disorder in the literature: alcoholism, substance abuse, panic disorder, psychosis, suicidal behavior, and thyroid disease (1). Thyroid disease data (hypo- or hyperthyroidism) were not collected during the first year of the study but were available from 312 affected and 149 unaffected subjects ascertained later.
Table 2. Clinical Features of Probands With Bipolar I, II, and Schizoaffective Bipolar Disorder and Their Relatives With Major Affective Disorder: Comparison Between Lowest and Highest Quartiles of Episode Frequency

<table>
<thead>
<tr>
<th>Clinical Feature</th>
<th>Subjects in the Lowest Quartile (N=98)</th>
<th>Subjects in the Highest Quartile (N=101)</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td>Episodes/year</td>
<td>0.16</td>
<td>0.06</td>
<td>0.02–0.28</td>
</tr>
<tr>
<td>Age at onset (years)</td>
<td>27.9</td>
<td>11.6</td>
<td>18.2–70</td>
</tr>
<tr>
<td><strong>Diagnosis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar I disorder</td>
<td>26</td>
<td>26.5</td>
<td>20</td>
</tr>
<tr>
<td>Bipolar II disorder</td>
<td>8</td>
<td>8.2</td>
<td>78</td>
</tr>
<tr>
<td>Recurrent unipolar depression</td>
<td>60</td>
<td>61.2</td>
<td>2</td>
</tr>
<tr>
<td>Schizoaffective disorder, bipolar type</td>
<td>4</td>
<td>4.1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>71</td>
<td>72.4</td>
<td>58</td>
</tr>
<tr>
<td>Male</td>
<td>27</td>
<td>27.6</td>
<td>43</td>
</tr>
<tr>
<td><strong>Axis I comorbidity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panic disorder</td>
<td>10</td>
<td>10.2</td>
<td>17</td>
</tr>
<tr>
<td>Alcoholism</td>
<td>21</td>
<td>21.4</td>
<td>43</td>
</tr>
<tr>
<td>Substance abuse/dependence</td>
<td>11</td>
<td>11.2</td>
<td>18</td>
</tr>
<tr>
<td>Suicidal behavior</td>
<td>6</td>
<td>6.1</td>
<td>21</td>
</tr>
</tbody>
</table>

* Significant overall difference between patients in the lowest and highest quartiles (χ²=113.7, df=1, p≤0.0001).

Statistical Analysis

Episode frequency was calculated on the basis of the total reported episodes of major depression, mania, or hypomania per year of illness, as follows: episode frequency = total number of episodes of illness/age at interview – age at onset. Five subjects with an illness duration of less than 2 years were excluded, because a small denominator may inflate the episode frequency and subjects may tend to volunteer for a study around a time of increased illness activity. Reports of mixed episodes are not elicited by the SADS-L, so data for mixed episodes are not included in this calculation. Because episode frequency showed a highly skewed distribution, the values were log-transformed.

The data were stored by using a relational database system that was based on Paradox (versions 5 and 8) (Corel Corp., Ottawa, Canada) (21).

Continuous variables were analyzed by t test or analysis of variance (ANOVA), and categorical variables were analyzed by Pearson’s chi-square test and Fisher’s exact test. Because the use of relatives’ data in the ANOVA could lead to biased estimates of variance, familiality was also assessed in a mixed-effects regression procedure in which the likelihood ratio test was used to compare the log likelihoods of “intercept-only” models and models that include family membership as a random effect (22). The mixed-effects model assumes that data within clusters are dependent (as is the case for data from relatives) and estimates the degree of dependency along with the parameters of the model.

The ANOVA was performed with SAS (SAS Institute, Cary, N.C.), and the regression analyses were performed with MIXREG (23). The alpha level was set at 0.05. All tests were two-tailed.

Results

Descriptive Statistics

The variables used in the analysis of episode frequency, along with their relationship to gender and diagnostic group, are shown in Table 1. There were no significant differences among diagnostic groups in mean age at onset, duration of illness, or gender, except for the expected higher prevalence of women in the recurrent unipolar depression group.

Episode frequency was smoothly distributed over the range of 0.02 to 20.2 episodes per year in this group of subjects (Figure 1). We observed no discontinuity at the DSM-IV rapid-cycling threshold of four episodes per year. The distributions were similar when depressive, manic, and hypomanic episodes were considered separately, so only the total episode frequency was considered in the subsequent analyses.

Association With Diagnostic Group

Episode frequency was significantly associated with diagnostic group (p<0.001). Specific comparisons revealed that episode frequency was highest among subjects with bipolar II disorder and lowest among subjects with recurrent unipolar depression (Bonferroni corrected p<0.001 for each of four comparisons). When subjects in the lowest and highest quartiles of episode frequency (range=0.02–0.28 episodes/year and range=2.3–20.2 episodes/year, respectively) were contrasted (Table 2), recurrent unipolar depression was more prevalent in the lowest quartile, and bipolar II disorder was more prevalent in the highest quartile (p≤0.0001).

Association With Other Variables

Gender. There was no significant difference in episode frequency between male and female subjects (t=1.39, df=403, n.s.), but subjects in the lowest quartile of episode frequency were more likely to be female (Table 2).

Age at onset. Age at onset was negatively correlated with episode frequency (r=-0.19, N=401, p<0.001). However, the relationship was not strictly linear (Figure 2). The highest episode frequencies were seen among subjects with onset between ages 15 and 18 years. Similarly, age at onset was significantly lower among those in the highest quartile of episode frequency (Table 2).
Psychosis. Psychosis, defined here as hallucinations or delusions (1), frequently complicated bipolar affective disorder in this group of subjects; 31.2% of subjects reported a history of psychosis at some time during their illness. Episode frequency was significantly higher in subjects with a history of psychosis than in those with no such history (F=5.53, df=1, 392, p<0.02).

Suicidal behavior and other axis I comorbidities. Typical patterns of suicidal behavior and comorbidity were observed in this group of subjects (1, 24, 25). Alcoholism was the most common comorbid disorder, found in 33.8% (N=136) of the subjects. Suicidal behavior was present in 20.6% (N=83), illicit substance abuse/dependence in 16.7% (N=67), and panic disorder in 11.9% (N=48).

Episode frequency was significantly associated with some but not all of these variables. Episode frequency was significantly associated with alcoholism (F=8.64, df=1, 400, p=0.003) and with a history of suicidal behavior (F=7.23, df=1, 400, p=0.007). Panic disorder and illicit substance abuse/dependence were not associated with episode frequency (F=2.49, df=1, 397, n.s., and F=2.35, df=1, 400, n.s., respectively). Similar results were observed in the quartiles analysis (Table 2), which also demonstrates the direction of each association.

Familiality

Episode frequency was significantly correlated among probands and their affected relatives (intraclass r=0.56, F=1.53, df=96, 321, p=0.004), suggesting that more than 30% of the variance in episode frequency was accounted for by family membership.

The familiality of episode frequency was confirmed in the mixed regression analysis (Table 3), which accounted for the nonindependence of data among relatives. Episode frequency remained strongly familial when the correlated variables (diagnostic group, age at interview, age at onset, psychosis, alcoholism, and suicidal behavior) were included in the model as fixed effects. This finding showed that the familiality of episode frequency in this group of subjects was not due solely to any familial tendency of the correlated variables.

Thyroid disease comorbidity

There was no significant difference in episode frequency between subjects with thyroid disorder and those without thyroid disorder. However, 43 (13.8%) of 312 affected subjects and nine (6%) of 149 unaffected subjects reported a history of thyroid disease (Table 4). A similar proportion of affected subjects who were exposed to lithium (19 [14.3%] of 132) and affected subjects who were not exposed to lithium (24 [13.3%] of 180) reported thyroid disease, and both of those groups were significantly more likely to report thyroid disease than were the unaffected subjects (p<0.05). The prevalence of thyroid disease is generally considered to be greater in females and to increase with age (26), but this pattern did not account for our findings. The proportion of female subjects was similar in the groups with and without thyroid disease, and the mean age at interview was lower among affected subjects with thyroid disease than among unaffected subjects (mean=46.6 years, SD=11.9, versus mean=68.5 years, SD=11.0) (t=3.9, df=30, p<0.001).

DSM-IV rapid cycling

Forty-six subjects (11.1%) met the DSM-IV criteria for rapid cycling (four or more episodes/year). These subjects were significantly more likely than those reporting fewer than four episodes/year to have a diagnosis of bipolar II disorder (87.1%) (χ²=55.2, df=1, p<0.0001). However, none of the other demographic or clinical variables analyzed in this study differed between the subjects with DSM-IV rapid cycling and the other subjects. We did not detect significant evidence that the categorical trait of rapid cycling was familial in this study group.

Discussion

To our knowledge, this study demonstrated for the first time that episode frequency, defined as the number of affective episodes per year, is a familial trait in bipolar affective disorder. We further showed that the familiality of episode frequency is not accounted for by other, correlated variables, such as affective diagnosis subtype, psychosis, alcoholism, or suicidal behavior. As a familial trait associated with several indicators of disease severity, epi-
sode frequency may help to define clinical subtypes of bipolar affective disorder with greater genetic liability. These data did not support DSM-IV rapid cycling as the best predictor of familiality or severity.

The data were collected from 407 subjects in 86 families, to our knowledge the largest data set ever studied for episode frequency. Diagnoses were highly reliable and were made by including all available data in a best-estimate procedure. The primary data used in estimating episode frequency were retrospective and thus subject to recall bias. There was no reason to expect, however, that relatives would be correlated in their recall bias; thus, recall bias alone cannot account for our findings. The data collection methods had other limitations. Subjects with two or more episodes of mania were not questioned in detail about all hypomanic episodes. Thus, true episode frequency was probably underestimated for subjects with bipolar I disorder. This underestimation was also unlikely to account for our main findings, because the exclusion of hypomanic episodes did not change the distribution of episode frequency in this study group. No data were collected on treatment, so it was not possible to control for potential treatment effects on episode frequency in these data.

DSM-IV defines subjects with four or more major affective episodes in a year as having rapid cycling. The published studies of rapid-cycling subjects, so defined, have not consistently detected familial aggregation (4, 5, 12–15). We found little support for the DSM-IV definition of rapid cycling in this study group. As a categorical trait, rapid cycling was not familial in this analysis, even though we found significant evidence of familiality when we considered subjects across the full range of episode frequency. Rapid cycling was associated with the diagnosis of bipolar II disorder, but it was not correlated with any of the other clinical features we examined. Our retrospective data did not allow us to identify subjects with discrete periods of rapid cycling punctuating a course of illness with few other episodes. We cannot rule out familial effects in these subjects.

The evidence linking hypothyroidism and rapid cycling is controversial. The association of hypothyroidism with major affective disorders and the presence of thyroid antibodies in patients with major affective disorders have been described in numerous publications, but the biological relationship between these two illnesses remains unclear (6–8, 10). Our data, which relied solely on subjects’ self-report, were limited by lack of information from thyroid function tests and thyroid antibodies assays. Nevertheless, the rate of reported thyroid disease among unaffected subjects (6%) appeared similar to the population rate of 5.9% (26), and the rate of reported thyroid disease among the lithium-treated subjects was similar to that reported in another study (27). Thus, we do not appear to have greatly under- or overestimated the rates of thyroid disease in these study subjects. We found, as have others, that thyroid disease was more common among affected subjects and that this difference remained significant even after subjects with a history of lithium exposure were dropped from the analysis. We found no association between thyroid disease and episode frequency in the study group, consistent with previous studies (4, 5, 28).

Alcoholism and substance abuse are known to be highly associated with major affective disorders (1, 24, 25). In a previous study that included some of the subjects also included in the present analysis, alcoholism was found to be associated with a higher rate of suicide attempts and to be clustered in a subset of families (29). The present study extended this finding by demonstrating that alcoholism and suicide attempts are both strongly associated with episode frequency. Our new results suggest that the association between alcoholism and suicidal behavior is mediated, at least in part, by an increase in episode frequency.
Episode frequency was also associated with other clinical features of bipolar affective disorder. Age at onset (itself a significant predictor of prognosis, comorbidity, and treatment response in bipolar affective disorder [30]) was strongly associated with episode frequency in this study group. We also found that episode frequency was significantly associated with psychotic features. It is possible that episode frequency accounts for some of the tendency of psychotic features to run in families with bipolar affective disorder (31, 32). An earlier analysis of a subset of these subjects indicated that episode frequency tended to increase in successive generations of a pedigree, a phenomenon known as anticipation (33). A complete analysis of anticipation is beyond the scope of the study reported here, but if anticipation were present in the current study group, it would tend to decrease the familiality of episode frequency and thus could not account for our findings.

It may come as a surprise that the highest quartile of episode frequency contained many subjects with bipolar II disorder, which is traditionally considered less severe than bipolar I disorder, as well as many subjects with early onset, psychotic features, alcoholism, and suicidal behavior, clear indicators of a severe illness. This finding is consistent with previous reports indicating more chronicity and a higher rate of comorbidity and suicidal behavior in bipolar II disorder, compared to bipolar I disorder (34, 35). These results imply that the traditional view of severity, which emphasizes mania, may be too narrow. Bipolar II disorder is in many ways more “severe” than bipolar I disorder.

These findings have implications for genetic research in bipolar affective disorder. As a quantitative trait, episode frequency may offer an alternative phenotype that would be more powerful than the categorical phenotypes typically used in genetic linkage and association studies. Episode frequency may also offer an approach to genetic heterogeneity in bipolar affective disorder, because subjects with similar episode frequencies may be more likely to share genetic determinants. Episode frequency is also associated with several indicators of disease severity. To the extent that disease severity is related to genetic liability, episode frequency may help to define clinical subtypes of bipolar affective disorder with a greater burden of genetic risk factors. However, further data concerning the heritability of episode frequency—for example, in twins—is needed before we can make confident predictions about the potential value of episode frequency as a phenotype in genetic research.

In conclusion, we found that episode frequency is a familial trait in bipolar affective disorder. These data also raise concerns about the DSM-IV definition of rapid cycling in bipolar affective disorder. We suggest that episode frequency is an important clinical feature of bipolar affective disorder, with implications for severity, comorbid conditions, and genetic research.

Presented in part at the 57th annual meeting of the Society of Biological Psychiatry, Philadelphia, May 16–18, 2002. Received Jan. 29, 2003; revision received June 24, 2004; accepted Aug. 2, 2004. From the Departments of Psychiatry and Medicine, University of Chicago, Chicago; the Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore; Division of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, University of Heidelberg, Mannheim, Germany; and the Unit on the Genetic Basis of Mood and Anxiety Disorders, NIMH, Bethesda, Md. Address correspondence and reprint requests to Dr. Fisfalen, Department of Psychiatry, Mount Sinai Medical Center, Rosalind Franklin University of Medicine and Science, California at 150th, Chicago, IL 60608; fisfalen@sinai.org (e-mail).

Supported by grants from the NIMH Intramural Research Program and NIH, the National Alliance for Research on Schizophrenia and Depression, the Chicago Brain Research Foundation, and the Edward F. Mallinckrodt, Jr., Foundation. Additional support for family recruitment was provided by the Charles A. Dana Foundation.

The authors thank Sylvia G. Simpson, Dean Mackinnon, and Melvin G. McInnis for contributing to the family evaluations, Donald Hede- ker for advice on the use of MIXREG, and the family volunteers who made this work possible.

References


Comparison of Rapid-Cycling and Non-Rapid-Cycling Bipolar Disorder Based on Prospective Mood Ratings in 539 Outpatients

Ralph W. Kupka, M.D., Ph.D.
David A. Luckenbaugh, M.A.
Robert M. Post, M.D.
Trisha Suppes, M.D., Ph.D.
Lori L. Altschuler, M.D.
Paul E. Keck, Jr., M.D.
Mark A. Frye, M.D.
Kirk D. Denicoff, M.D.
Heinz Grunze, M.D.
Gabriele S. Leverich, M.S.W.
Susan L. McElroy, M.D.
Jörg Walden, M.D., Ph.D.
Willem A. Nolen, M.D., Ph.D.

Objective: To detect risk factors for rapid cycling in bipolar disorder, the authors compared characteristics of rapid-cycling and non-rapid-cycling patients both from a categorical and a dimensional perspective.

Method: Outpatients with bipolar I disorder (N=419), bipolar II disorder (N=104), and bipolar disorder not otherwise specified (N=16) were prospectively evaluated with daily mood ratings for 1 year. Subjects were classified as having rapid cycling (defined by the DSM-IV criterion of four or more manic or depressive episodes within 1 year) or not having rapid cycling, and the two groups’ demographic and retrospective illness characteristics were compared. Associated factors were also evaluated in relationship to episode frequency.

Results: Patients with rapid cycling (N=206; 38.2%) significantly differed from those without rapid cycling (N=333) with respect to the following independent variables: history of childhood physical and/or sexual abuse, bipolar I disorder subtype, number of lifetime manic or depressive episodes, history of rapid cycling, and history of drug abuse. The prevalence of these characteristics increased progressively with episode frequency. The proportion of women was greater than the proportion of men only among patients with eight or more episodes per year. The average time spent manic/hypomanic increased as a function of episode frequency, but the average time spent depressed was comparable in patients with one episode and in those with more than one episode. Brief episodes were as frequent as full-duration DSM-IV-defined episodes.

Conclusions: A number of heterogeneous risk factors were progressively associated with increasing episode frequency. Depression predominated in all bipolar disorder patients, but patients with rapid cycling were more likely to be characterized by manic features. The findings overall suggest that rapid cycling is a dimensional course specifier arbitrarily defined on a continuum of episode frequency.

In 1974, Dunner and Fieve (1) defined patients with rapid cycling as having four or more manic or depressive episodes of at least 2 weeks’ duration in the year before the study. Although subsequent studies addressed the issue of rapid cycling from the perspectives of clinical and demographic features, neurobiological dysfunction, longitudinal course, and treatment response, it remains unclear to what extent rapid cycling delineates a distinct subtype or is merely an arbitrary point on a continuum of episode frequencies.

The validity of rapid cycling as a course specifier for DSM-IV bipolar I disorder and bipolar II disorder was reviewed by Bauer and Whybrow (2). They supported the inclusion of rapid cycling as a course specifier despite important unresolved issues, such as the paucity of characteristics that distinguish rapid-cycling patients (only female sex emerged from most studies), the arbitrariness of the criterion of four or more mood episodes per year, and the occurrence of brief but severe episodes that appeared to be abundant in rapid cycling but were formally excluded by the DSM-IV criteria. Subsequent studies gave further support for the inclusion of brief episodes in the criteria but provided little evidence in support of the boundary of four episodes per year (3–6).

In a meta-analysis of 20 clinical studies comparing rapid-cycling and non-rapid-cycling bipolar disorder (6), the overall prevalence of rapid cycling in unselected research samples was 16.3%, and rapid cycling was significantly associated with female gender and the bipolar II disorder subtype. That analysis confirmed the need for prospective evaluation in a large sample, particularly to address interrelationships among risk factors (3, 7).

To address these issues, we assessed episode frequency from prospective daily mood ratings and compared retrospective and prospective illness variables in patients with a rapid-cycling and a non-rapid-cycling course. In addition to classifying rapid cycling by using the traditional definition of four or more mood episodes per year, we examined these clinical characteristics in relationship to the dimension of episode frequency in an attempt to dis-
cern an empirically based boundary for rapid cycling, if present.

Method

Subjects

Subjects were adult outpatients with DSM-IV bipolar I disorder (N=419, 77.7%), bipolar II disorder (N=104, 19.3%), and bipolar disorder not otherwise specified (N=16, 3.0%) diagnosed with the Structured Clinical Interview for DSM-IV (SCID) (8). All patients participated in the former Stanley Foundation Bipolar Network with four sites in the United States (N=88, N=104, N=86, N=73), one in the Netherlands (N=143), and two in Germany (N=28, N=17), and were enrolled from 1995 through 2000 (9). Outpatients were recruited from the participating clinics and in many cases were referred by local physicians or were self-referred from local advocacy groups; the only exclusion criterion was a current comorbid substance use disorder that was severe enough to require treatment in a specialized setting (10). Interrater reliability for the diagnosis of bipolar disorder was excellent (overall kappa of 0.92) (11).

Patients received treatment according to current standards in a naturalistic fashion or in various more formal pharmacological treatment protocols also designed to match usual treatment. Patients in this study group may have been included in studies or clinical trials described in other published reports from the Stanley Foundation Bipolar Network. The study was approved by the institutional review boards of all participating sites, and all patients provided written informed consent.

Demographic and Clinical Characteristics

Patient- and clinician-rated questionnaires were used to gather data on demographic characteristics, functional status, prior illness characteristics, and psychiatric history of the parents as discussed elsewhere (9–12). The SCID was used to determine co-morbid axis I diagnoses. In addition, measures of thyroid parameters were available for a subgroup of 199 patients who simultaneously participated in another study (13).

Prospective Follow-Up and Episode Frequency

The prospective course of illness was followed with the clinician-rated NIMH Life-Chart Method (14), which is used to record daily ratings of severity of manic and depressive symptoms on a 9-point graphic scale. Severity ratings were based on symptom-driven degree of functional impairment. At the baseline level (eu- thymia, score 0), there were neither significant mood symptoms nor functional impairment. Ratings for mania were as follows: 2.5=mild, 5=low moderate, 7.5=high moderate, and 10=severe. Depression ratings were as follows: ≥2.5=mild, ≤5=low moderate, ≤7.5=high moderate, and ≤10=severe. A patient-rated Life-Chart Method was evaluated weekly to monthly by a research clinician, the various mood states and determined the mean level of severity and duration: 4 days of mild ratings for hypomania, 1 week of moderate ratings or any hospitalization for mania, and 2 weeks of moderate ratings for depression. An episode terminated with any switch to the opposite polarity or after 2 months of euthymia. By using a second set of criteria, we also identified isolated severe episodes that lasted 1 or more days. The computer program calculated the number of days the patient experienced the various mood states and determined the mean level of severity of the mood episodes. Computer-calculated episode counts based on the DSM-IV criteria were compared to visually counted episodes from printed life charts in a subsample of 63 patients. The correlation between the two measures was significant for depressive episodes (r=0.77, p<0.0001) and for hypomanic, manic, and mixed episodes (r=0.99, p<0.0001).

Data for all patients with at least 1 year of uninterrupted prospective daily Life-Chart Method ratings were included in the analysis. We excluded patients who left the Network before 1 year and patients with missing daily ratings at the time of the analysis. Patients with rapid cycling were defined as those with four or more depressive, hypomanic, manic, or mixed episodes (according to the DSM-IV duration criteria) in the first year of prospective follow-up.

Data Analysis

We compared demographic and retrospective and prospective illness characteristics of patients with DSM-IV-defined rapid cycling and patients without rapid cycling. We used chi-square analyses with Yates’s corrections for dichotomous variables and independent-sample t tests for continuous variables and applied Hochberg’s adjusted Bonferroni procedure for multiple tests of significance (17). For the multivariate analysis, logistic regression was used to examine significant independent contributions of the risk factors to rapid cycling. Variables that were significant in the univariate analyses were included in the logistic regression analysis. For risk factors that were highly intercorrelated, only one of the two factors was included in the model. Other risk factors were removed if multicollinearity was an issue in any given model.

Variables that distinguished between patients with and without rapid cycling were also evaluated for their relationship with episode frequency in the first year of prospective follow-up to reveal discontinuities indicating a potential optimal cutoff point for rapid cycling. Pearson’s product-moment correlations were used to supplement visual inspection of the curves for nonlinear relationships. All statistical analyses were performed with SPSS Version 9.0 (SPSS, Inc., Chicago).

Results

Comparison of Patients With and Without Rapid Cycling

The 539 patients with 1 year of prospective ratings included 237 men (44.0%) and 302 women (56.0%). Their mean age was 42.1 years (SD=11.5). A total of 206 (38.2%) had rapid cycling, and 333 (61.8%) did not have rapid cycling. A lifetime history of rapid cycling was reported at study entry by 264 (50.7%) of 521 patients. The demographic and clinical characteristics of the current study group did not differ from those of a larger group of 631 patients that included patients with a shorter follow-up period or incomplete ratings (12).

The patients’ clinical characteristics at study entry are summarized in Table 1. There was a nonsignificant over-representation of women in the rapid-cycling group. Educational level, marital status, and work status did not differ significantly between the groups (data not shown). Rapid cycling occurred in 41.3% of patients with bipolar I disorder and 27.9% of patients with bipolar II disorder. Patients with rapid cycling had an earlier age at onset, a longer duration of illness, and a longer time from first symptoms to first medication treatment, compared to the patients with-
out rapid cycling. There was a strong relationship between the occurrence of rapid cycling in any prior year, in the year before study entry, and during prospective follow-up. Rapid cycling was associated with a prior history of dysphoric mania/hypomania, lifetime treatment with antidepressants, and a history of substance-induced episodes.

Rapid cycling was associated with lifetime DSM-IV anxiety disorder, childhood physical or sexual abuse, and parental history of drug abuse, as detailed in Table 2.

Multivariate logistic regression (N=397) identified five independent variables that were significantly associated with rapid cycling: a lifetime history of rapid cycling (Wald's $\chi^2=17.64$, df=1, $p<0.0001$, odds ratio=3.39), higher number of lifetime mood episodes at study entry (11–20 episodes: Wald's $\chi^2=5.24$, df=1, $p=0.02$, odds ratio=4.66; >20 episodes: Wald's $\chi^2=8.51$, df=1, $p=0.004$, odds ratio=6.49), bipolar I disorder subtype (Wald's $\chi^2=8.22$, df=1, $p=0.004$, odds ratio=2.76), lifetime history of drug abuse (Wald's $\chi^2=4.57$, df=1, $p=0.03$, odds ratio=1.99), and history of childhood physical and/or sexual abuse (Wald's $\chi^2=4.67$, df=1, $p=0.03$, odds ratio=1.86).

Prospective Course of Illness

As summarized in Table 3, patients with rapid cycling had a sevenfold greater mean number of full-duration manic and hypomanic episodes and a twofold greater mean number of depressive episodes, compared to patients without rapid cycling. Over the course of the year, patients with rapid cycling spent many more days in a manic or hypomanic state than did patients without rapid cycling (73 days versus 25 days). Patients with rapid cycling also spent more days depressed, although the differ-

---

TABLE 1. Characteristics of Bipolar Disorder Patients With and Without Rapid Cycling in a 1-Year Prospective Study of Illness Course

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Patients With Rapid Cycling (N=206)</th>
<th>Patients Without Rapid Cycling (N=333)</th>
<th>Analysisc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At study entry</td>
<td>42.4 11.1</td>
<td>42.5 11.9</td>
<td>0.11 537 0.91</td>
</tr>
<tr>
<td>At onset of first mood symptoms that affected functioning</td>
<td>17.6 9.1</td>
<td>23.1 10.0</td>
<td>5.98 462 &lt;0.0001d</td>
</tr>
<tr>
<td>At first depressive symptoms</td>
<td>18.9 9.9</td>
<td>23.9 10.5</td>
<td>5.17 462 &lt;0.0001d</td>
</tr>
<tr>
<td>At first manic/hypomanic symptoms</td>
<td>21.7 10.1</td>
<td>27.2 11.3</td>
<td>5.36 462 &lt;0.0001d</td>
</tr>
<tr>
<td>Time from first symptoms to first medication treatment (years)</td>
<td>11.2 10.3</td>
<td>6.8 8.9</td>
<td>4.73 450 &lt;0.0001d</td>
</tr>
<tr>
<td>Duration of illness from first mood symptoms (years)</td>
<td>24.3 11.8</td>
<td>19.4 11.6</td>
<td>4.39 462 &lt;0.0001d</td>
</tr>
</tbody>
</table>

Global Assessment of Functioning Scale score

| Past week before study entry                             | 61.4 13.9                          | 67.5 15.2                            | 4.41 495 <0.0001d |
| Best week in year before study entry                     | 74.4 12.1                          | 77.8 11.9                           | 2.92 459 0.004d |
| Worst week in year before study entry                    | 43.5 15.1                          | 48.7 17.2                           | 3.24 457 0.001d |

<table>
<thead>
<tr>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender</td>
<td>129 62.6</td>
<td>173 52.0</td>
<td>5.46 1 0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder subtype</td>
<td></td>
<td></td>
<td>7.56 2 &lt;0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar I disorder</td>
<td>173 84.0</td>
<td>246 73.9</td>
<td>515 &lt;0.0001d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar II disorder</td>
<td>29 14.1</td>
<td>75 22.5</td>
<td>515 &lt;0.0001d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder not otherwise specified</td>
<td>4 1.9</td>
<td>12 3.6</td>
<td>515 &lt;0.0001d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifetime number of mood episodes at study entry</td>
<td>66.12 94.0</td>
<td>2.02 5.9</td>
<td>515 &lt;0.0001d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–4</td>
<td>5 2.6</td>
<td>38 12.8</td>
<td>515 &lt;0.0001d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5–10</td>
<td>12 6.2</td>
<td>75 25.3</td>
<td>515 &lt;0.0001d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11–20</td>
<td>25 13.0</td>
<td>56 18.9</td>
<td>515 &lt;0.0001d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;20</td>
<td>151 78.2</td>
<td>127 45.7</td>
<td>515 &lt;0.0001d</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rapid-cycling status

| Lifetime history                                        | 147 74.6                           | 117 36.1                            | 71.15 1 <0.0001d |
| In last year before study entry                         | 99 58.2                           | 63 27.5                            | 36.93 1 <0.0001d |
| Psychotic mood episodes: lifetime history               | 114 61.3                           | 173 55.6                            | 1.31 1 0.25 |
| Dysphoric mania/hypomania: lifetime history             | 137 71.7                           | 153 48.1                            | 26.19 1 <0.0001d |
| One or more serious suicide attempts: lifetime history  | 70 35.7                           | 82 26.8                            | 4.09 1 <0.05 |
| Exposure to antidepressants: lifetime history           | 187 93.5                           | 268 83.8                            | 9.82 1 0.002d |
| Antidepressant-induced mania/hypomania: lifetime history| 107 61.1                           | 118 49.0                            | 5.58 1 0.018 |
| Alcohol- or drug-induced mood episode: lifetime history | 69 35.8                           | 63 21.0                            | 12.29 1 <0.0001d |

Thyroid status

| Lifetime history of hypothyroidismf                   | 48 25.5                           | 55 18.3                            | 3.18 1 0.08 |
| Positive thyroid autoantibody testg                    | 30 33.3                           | 27 24.8                            | 1.37 1 0.24 |

a Number of subjects may differ slightly among variables because of missing data.
b Rapid cycling was defined as four or more DSM-IV mood episodes recorded with the NIMH Life-Chart Method during the 1-year prospective follow-up.
c Yates’s correction was used for chi-square tests for all two-by-two comparisons.
d Statistically significant after Hochberg’s adjusted Bonferroni procedure for multiple tests.

e In subjects with a lifetime exposure to antidepressants.
f Most cases of clinical and subclinical hypothyroidism were currently or previously treated with thyroid hormone.
g Data from 199 subjects also included in a previous study (13).
ence between groups was smaller (145 days versus 121 days). During follow-up, patients with rapid cycling were twice as likely to experience dysphoric mania/hypomania as were patients without rapid cycling (82.5% versus 38.1%) ($\chi^2=9.56$, df=1, $p<0.0001$).

Lithium, valproate, carbamazepine, or lamotrigine was used alone or in combination by 193 (93.7%) of the patients with rapid cycling and 305 (91.6%) of the patients without rapid cycling. Patients with rapid cycling were more likely than those without rapid cycling to be given antidepressants (67% versus 57%) ($\chi^2=4.60$, df=1, $p=0.03$), antipsychotics (52% versus 32%) ($\chi^2=19.08$, df=1, $p<0.0001$), and thyroid hormone (32% versus 22%) ($\chi^2=6.71$, df=1, $p=0.01$).

### Variables Related to Episode Frequency

The subjects were grouped by the number of full-duration DSM-IV episodes in the first year of prospective follow-up, as follows: no episodes (N=74) and one (N=110), two (N=78), three (N=71), four (N=46), five (N=39), six (N=24), seven (N=24), eight (N=22), nine (N=13), and 10 or more episodes (N=38).

Figure 1 illustrates prospective illness characteristics in relation to overall episode frequency. Increasing episode frequency was largely attributable to manic and hypomanic episodes. The number of DSM-IV depressive episodes was relatively constant in the groups of patients with five or more full-duration DSM-IV-defined mood episodes per year. Similarly, the average number of days per year during which patients experienced hypomanic, manic, or ultradian cycling gradually increased with the number of full-duration DSM-IV-defined episodes per year, and the average number of days during which patients were depressed was relatively consistent among the groups of patients with one or more episodes per year. The number of additional brief episodes progressively increased along with the frequency of full-duration DSM-IV-defined episodes.

Figure 2 shows the relationship between the overall frequency of full-duration DSM-IV-defined episodes and the prevalence of the variables that distinguished patients with and without rapid cycling. Except for bipolar I disorder subtype and female gender, these variables tended to increase in prevalence as a function of episode number ($r=0.15-0.37$, all $p<0.0001$). The prevalence of bipolar I disorder subtype had a somewhat uneven pattern. An increase in the percentage of female patients at eight episodes could indicate nonlinearity; 53% of patients with zero to seven episodes were women, compared to 73% of patients with eight or more episodes ($\chi^2=8.65$, df=1, $p=0.003$), although the percentage of women among patients with nine and 10 or more episodes was lower than the percentage of women among patients with eight episodes. None of the other curves showed evidence of nonlinearity suggestive of this or any other cutoff point for separating patients with rapid cycling from those without rapid cycling. We reanalyzed the comparisons reported in Table 1 and Table 2 with a definition of rapid cycling as eight or more episodes; 73 subjects had rapid cycling according to this definition. Female gender was the only variable with a greater significance in the reanalysis than in the original analysis (data not shown).

### Discussion

To our knowledge, this study is the first comparative investigation of rapid-cycling and non-rapid-cycling bipolar disorder that systematically used both a categorical and a dimensional analysis and is one of the few studies (18, 19)
Patients with rapid cycling were depressed 39.5% of the time, major of patients with bipolar illness is depression (21); patients with rapid cycling were considerably greater severity of depressive episodes, compared to patients without rapid cycling. Patients with rapid cycling had as many brief mood episodes as full-duration DSM-IV episodes. However, there was a lack of clear boundaries between the patients with and without rapid cycling on any of the prospective and retrospective variables examined as a function of episode number.

The prospective course of illness indicates that the study subjects were considerably ill despite extensive treatment. The prospective illness characteristics of the 419 bipolar I disorder patients were similar to findings in a group of 146 patients with bipolar I disorder who were followed for 2–20 years in the Collaborative Depression Study (20). The patients in the Collaborative Depression Study were depressed an average of 31.9% of the time, manic for 2.3%, hypomanic for 7.0%, and cycling/mixed for 5.9%. The patients in our study experienced those mood states 35.6%, 3.9%, 8.7%, and 3.3% of the time, respectively.

These findings confirm that the main burden for a majority of patients with bipolar illness is depression (21); patients with rapid cycling were depressed 39.5% of the time, and patients without rapid cycling were depressed 33.2% of the time. It is interesting to note that in all patients (both with and without rapid cycling) who had one or more full-duration episodes per year, the average amount of time in a depressed state was fairly stable, regardless of the total number of episodes (Figure 1). In contrast, the proportion of time in a hypomanic, manic, or ultradian cycling state was significantly higher in patients with rapid cycling (27.1% of the time versus 7.7% of the time in patients without rapid cycling) and increased progressively as a function of episode frequency.

When we compared the prior illness histories of the patients with and without rapid cycling, we found that five independent variables were associated with rapid cycling in the logistic regression analysis: previous rapid cycling, a greater number of previous mood episodes, bipolar I disorder subtype, history of childhood physical and/or sexual abuse, and lifetime drug abuse.

A previous history of rapid cycling and a history of more than 10 mood episodes before study entry were the strongest predictors of rapid cycling during the first year of prospective follow-up. This finding suggests that some patients have a propensity toward rapid cycling that may be expressed periodically or more continuously.

Our finding that rapid cycling was more prevalent among patients with bipolar I disorder contrasts with various reports of overrepresentation of patients with rapid cycling.
among bipolar II disorder patients (6). However, two large studies also failed to report a preponderance of rapid-cycling patients among patients with bipolar II disorder (22, 23).

Childhood physical or sexual abuse was associated with higher episode frequency. We previously reported that early abuse was associated with an earlier age at onset of bipolar disorder, serious suicide attempts, and comorbid drug abuse and anxiety disorders (12). These findings suggest that early traumatic experiences may contribute to a later adverse illness course.

Substance abuse is highly comorbid with bipolar disorder (11), and patients with rapid cycling may be particularly sensitive to the destabilizing properties of alcohol and drugs, as they frequently reported prior induction of depressive and manic/hypomanic episodes by these agents.
Alternatively, patients with faster cycling frequencies may use these substances at a higher rate.

Several factors were associated with rapid cycling in the univariate but not in the multivariate analysis, indicating that these factors were intercorrelated with one or more of the factors mentioned earlier. As in other studies (6), we found a modest overrepresentation of women among the patients with rapid cycling, especially among those with higher episode frequencies. Our finding of a younger age at onset of bipolar disorder in patients with rapid cycling was also found in previous studies (23–25), as was our finding of a longer duration of illness in patients with rapid cycling (4, 5), although most studies reviewed found no differences (6). Our finding of a longer time between first symptoms and first medication treatment in patients with rapid cycling raises the question of whether earlier recognition and treatment might prevent or attenuate this more problematic course.

Prior histories of both dysphoric mania/hypomania and ultradian cycling were more prevalent in patients with rapid cycling than in patients without rapid cycling. These associations were confirmed during prospective follow-up, which suggests that depressive features pervade both manic and depressive phases of the illness in patients with faster cycle frequencies. Our study was not designed to evaluate the contribution of antidepressants to the development of rapid cycling. Still, 54% of patients in the study who had been treated with antidepressants reported having experienced an antidepressant-induced switch to mania/hypomania in the past, a phenomenon that has been associated with a subsequent rapid-cycling course (26).

The most rigorous studies of family history of mood disorders reported no significant differences in family history between patients with and without rapid cycling (4–6, 19, 27, 28). We found that a parental history of major mood disorder (bipolar disorder and/or depression) was more likely in patients with rapid cycling than in those without rapid cycling. More prominent was our finding of a parental history of drug abuse in patients with rapid cycling. However, rapid cycling was also associated with the occurrence of childhood adversity, which suggests that parental substance abuse and the associated environmental instability may have indirectly contributed to the risk of rapid cycling.

The results of our study must be considered in the context of several limitations. Although our study population is in many respects comparable to other groups of outpatients (10), 80% of the subjects in our study had bipolar I disorder, almost 60% reported a lifetime history of psychotic symptoms, and 38% had a rapid-cycling course. This greater overall severity of illness may restrict the generalizability of our findings. The exclusion of patients with less than 1 year of follow-up may have affected the representativeness of the study subjects. Moreover, the retrospective data on previous illness characteristics covered a period of many years and thus were susceptible to inaccu-

Presented at the Fifth International Conference on Bipolar Disorder, Pittsburgh, June 12–14, 2003. Received Feb. 19, 2004; revisions received May 14 and June 29, 2004; accepted Aug. 2, 2004. From Albert Institute for Mental Health Care and University Medical Center Utrecht, Utrecht, Netherlands; the Mood and Anxiety Disorders Program and the Biological Psychiatry Branch, NIMH, Bethesda, Md.; the University of Texas Southwestern Medical Center, Dallas; the UCLA Mood Disorders Research Program, Los Angeles; the Psychopharmacology Research Program, Department of Psychiatry, University of Cincinnati, Cincinnati, Ohio; the Department of Psychiatry, University of Groningen, Groningen, Netherlands. Address correspondences to Dr. Kucpa, Albert Institute for Mental Health Care, Tolsteegsingel 2A, 3582 AC Utrecht, Netherlands; r.kupka@planet.nl (e-mail).

Supported by the Stanley Mental Research Institute.

The authors thank Adriaan Honig, M.D., Ph.D., Baer Arts, M.D., Titus van Os, M.D., Ph.D., Pieterem Kolling, M.D., Herro Kraan, M.D., Ph.D., Max Sonnen, M.D., Rocco Hoeckstra, M.D., Onno Habekotté, M.D., Harm-Jan Pot, M.D., Albert Blom, M.D., and Rein Holleboom, M.D., for patient recruitment and data collection in the Netherlands.

References


Am J Psychiatry 162:7, July 2005
http://ajp.psychiatryonline.org

KUPKA, LUCKENBAUGH, POST, ET AL.
RAPID-CYCLING BIPOLAR DISORDER


Olanzapine Versus Lithium in the Maintenance Treatment of Bipolar Disorder: A 12-Month, Randomized, Double-Blind, Controlled Clinical Trial

Mauricio Tohen, M.D., Dr.P.H.
Waldemar Greil, M.D.
Joseph R. Calabrese, M.D.
Gary S. Sachs, M.D.
Lakshmi N. Yatham, M.B., F.R.C.P.C.
Bruno Müller Oerlinghausen, Dr.Med.
Athanasios Koukopoulos, M.D.
Giovanni B. Cassano, M.D.
Heinz Grunze, M.D.
Rasmus W. Licht, M.D., Ph.D.
Liliana Dell’Osso, M.D.
Angela R. Evans, Ph.D.
Richard Risser, M.Sc.
Robert W. Baker, M.D.
Heidi Crane, M.S.
Martin R. Dossenbach, M.D.
Charles L. Bowden, M.D.

Objective: The authors compared the efficacy of olanzapine and lithium in the prevention of mood episode relapse/recurrence.

Method: Patients with a diagnosis of bipolar disorder (manic/mixed), a history of two or more manic or mixed episodes within 6 years, and a Young Mania Rating Scale total score ≥20 entered the study and received open-label cotreatment with olanzapine and lithium for 6–12 weeks. Those meeting symptomatic remission criteria (Young Mania Rating Scale score ≤12; 21-item Hamilton depression scale score ≤8) were randomly assigned to 52 weeks of double-blind monotherapy with olanzapine, 5–20 mg/day (N=217), or lithium (target blood level: 0.6–1.2 meq/liter) (N=214).

Results: Symptomatic relapse/recurrence (score ≥15 on either the Young Mania Rating Scale or Hamilton depression scale) occurred in 30.0% of olanzapine-treated and 38.8% of lithium-treated patients. The noninferiority of olanzapine relative to lithium (primary objective) in preventing relapse/recurrence was met, since the lower limit of the 95% confidence interval on the 8.8% risk difference (–0.1% to 17.8%) exceeded the predefined noninferiority margin (–7.3%). Secondary results showed that compared with lithium, olanzapine had significantly lower risks of manic episode and mixed episode relapse/recurrence. Depression relapse/recurrence occurred in 15.7% of olanzapine-treated and 10.7% of lithium-treated patients. Mean weight gain during open-label cotreatment was 2.7 kg; during double-blind monotherapy, weight gain was significantly greater with olanzapine (1.8 kg) than with lithium (–1.4 kg).

Conclusions: These results suggest that olanzapine was significantly more effective than lithium in preventing manic and mixed episode relapse/recurrence. Both agents were comparable in preventing depression relapse/recurrence.

Despite ongoing maintenance therapy, patients with bipolar disorder will experience frequent fluctuations in symptom severity and multiple relapses. Prospective naturalistic studies examining relapse have reported relapse risks ranging from 44% in 1 year (1) to 73%–88.7% over 4–5 years (2, 3). To date, a limited number of therapeutic agents have been available for the long-term treatment of bipolar disorder. Lithium has been the mainstay of maintenance therapy for >30 years. It is the most extensively studied mood stabilizer and has the best overall efficacy for the prophylactic treatment of bipolar disorder. Recent meta-analyses show it to be superior to placebo in the prevention of relapse (4) and to reduce the risk of relapse 3.6-fold (5). The anticonvulsant agents valproate (6), lamotrigine (7), and carbamazepine (8) also have been used as maintenance therapies.

Olanzapine has shown superiority to placebo in treating acute manic episodes in patients with bipolar I disorder (9, 10). Furthermore, in a 47-week comparative trial of olanzapine versus valproate, rates of mood episode following acute remission of mania were comparable between the groups (11), suggesting that olanzapine may be effective in the prevention of bipolar disorder relapse/recurrence.

This trial compared the efficacy of olanzapine and lithium for the prevention of mood episode relapse/recurrence. For simplicity, the term recurrence will be used throughout the text.
OLANZAPINE VERSUS LITHIUM IN BIPOLAR MAINTENANCE

TABLE 1. A Priori Categorical Definitions of Remission During Olanzapine and Lithium Cotreatment and Recurrence Following Random Assignment to Double-Blind Olanzapine or Lithium Monotherapy in Patients With Bipolar Disorder

<table>
<thead>
<tr>
<th>Categorical Definition</th>
<th>Remission</th>
<th>Recurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptomatic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young Mania Rating Scale total ≤12; and Hamilton depression scale score ≤8</td>
<td>Score ≥15 on Young Mania Rating Scale and/or Hamilton depression scale</td>
<td></td>
</tr>
<tr>
<td>Mania: All DSM-IV A and B criteria for current manic episode no worse than mild (≤3 on a scale of 1–7) and no more than two B criteria given mild rating (3 on a scale of 1–7)</td>
<td>Meeting DSM-IV criteria for current manic, depressive, or mixed episode</td>
<td></td>
</tr>
<tr>
<td>Depression: All DSM-IV A criteria for current major depressive episode no worse than mild (≤3 on a scale of 1–7) and no more than three A criteria given mild rating (3 on a scale of 1–7)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Syndromic

<table>
<thead>
<tr>
<th>Categorical Definition</th>
<th>Remission</th>
<th>Recurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mania: All DSM-IV A and B criteria for current manic episode no worse than mild (≤3 on a scale of 1–7) and no more than two B criteria given mild rating (3 on a scale of 1–7)</td>
<td>Meeting DSM-IV criteria for current manic, depressive, or mixed episode</td>
<td></td>
</tr>
</tbody>
</table>

3 Not including duration criterion.

Method

Patients

Patients enrolled in this study were ≥18 years of age and met DSM-IV criteria for bipolar disorder (current episode manic or mixed) as determined with the Structured Clinical Interview for DSM-IV, Patient Version. Patients were required to have a Young Mania Rating Scale total score ≥220 at baseline and a history of at least two manic or mixed episodes in the preceding 6 years. Patients were excluded from the study if they had a serious, unstable medical illness; met DSM-IV substance dependence criteria (nicotine or caffeine excepted) within the past 30 days; had been treated with a depot neuroleptic within 6 weeks of random assignment; or were considered a serious suicide risk. Patients were also excluded if they had a history of intolerance, or lack of response, to an adequate trial of lithium or olanzapine as determined by the investigator. After the study was completely described to the patients, written informed consent was obtained. The study was approved by the appropriate ethics review boards.

Study Design

Patients were recruited from 87 inpatient and outpatient settings across Western Europe, Canada, South Africa, Israel, Australia, and New Zealand between August 1999 and June 2002. This randomized, double-blind, controlled trial consisted of four study periods: 1) screening (two clinic visits over 2–7 days), 2) open-label cotreatment (6–12 weeks; twice-weekly visits for the first 2 weeks, weekly thereafter), 3) double-blind taper (4 weeks; weekly visits), and 4) double-blind monotherapy (48 weeks; biweekly during the first 4 weeks, monthly thereafter). Eligible patients began open-label cotreatment with olanzapine, 15 mg/day, and lithium, 600 mg/day. Allowed dosages of olanzapine were 5–20 mg/day. Investigators were required to optimize lithium dose and reach a target blood level of 0.6–1.2 mg/liter by week 4 during this period.

Patients who met symptomatic remission criteria (Table 1) during the open-label cotreatment period were randomly reassigned in a 1:1 ratio by means of a unique drug kit number (via a call-in Interactive Voice Response System) to monotherapy with either olanzapine or lithium. All patients, study site personnel, and sponsor investigators were blind to randomization codes. During the double-blind taper period, patients remained on their current dose of randomly assigned treatment, and the dose of the discontinued drug was tapered in a blinded, a priori-determined, stepwise manner over 4 weeks.

Lithium levels were monitored every 2 weeks during the double-blind taper period and monthly during double-blind maintenance monotherapy. If the serum level of lithium deviated from the therapeutic range during these study periods, the investigator was to adjust the dose of lithium to reestablish blood levels within the therapeutic range, with a goal of reaching this range within 30 days. Serum levels ranging from 0.6–1.2 mg/liter were considered within normal limits. Maintenance of the blind associated with blood draws has been described (12). Briefly, all patients randomly assigned to olanzapine also had blood drawn. For every outlier report generated for a lithium patient, a sham lithium outlier report was sent to an olanzapine patient. Thus, reports to investigative sites indicating that the lithium dose should be adjusted did not unmask the olanzapine patient.

Concomitant Medications

Patients who entered the study receiving psychotropic medications (including anticonvulsants, typical or atypical antipsychotics [oral or intramuscular], or antidepressants) were gradually discontinued from these medications at the discretion of the investigator during the first 3 weeks of the open-label cotreatment period. However, oral or intramuscular haloperidol and zuclopenthixol were permitted for extreme agitation during the open-label period. Benzodiazepines were allowed according to the following guidelines. The maximum dose from the screening period through the first 6 weeks of the open-label cotreatment period was 8 mg/day in lorazepam equivalents and 6 mg/day in lorazepam equivalents for the remainder of the open-label period and during the first 2 weeks of the taper period. It was further decreased to 4 mg/day for the remaining 2 weeks of the taper period and then to 2 mg/day (for not more than 60 cumulative days) for the double-blind monotherapy period. Patients were permitted concomitant medication for treatment-emergent extrapyramidal symptoms (biperiden or benzotropine mesylate, ≤8 mg/day; trihexyphenidyl, ≤12 mg/day). However, prophylactic use of anticholinergics for extrapyramidal symptoms was not allowed.

Assessments

Recurrence and severity of illness were assessed with the Young Mania Rating Scale and the 21-item Hamilton depression scale; raters were trained and certified to use these scales. A minimum reliability score (intraclass correlation [ICC]) of 0.75 was required for certification; raters who failed to achieve an ICC of ≥0.75 were retrained and tested again. Four hundred fifty-two raters were certified to use the Young Mania Rating Scale and Hamilton depression scale for this study, with average ICCs of 0.88 (Hamilton depression scale) and 0.90 (Young Mania Rating Scale). The vast majority of raters achieved an ICC of ≥0.85 (78.8% for the Hamilton depression scale and 77.4% for the Young Mania Rating Scale). Patient safety was evaluated through standard clinical observations, and extrapyramidal symptoms were assessed with the Simpson-Angus Rating Scale, the Barnes Rating Scale for Drug-Induced Akathisia, and the Abnormal Involuntary Movement Scale. Criteria for treatment-emergent extrapyramidal symptoms have been described previously (12).

Statistical Methods

The primary objective of the study was the assessment of olanzapine's noninferiority to lithium in the risk of symptomatic mood episode recurrence. It was estimated that it would require that 200 patients in symptomatic remission be randomly assigned to each therapy to provide 80% power to detect the protocol-defined mar-
One hundred seventy-one patients completed the double-blind phase of the study also shows significantly higher completion with sustained remission for olanzapine (43.3% [N=94 of 217]) than lithium (28.5% [N=61 of 214]) (p=0.002, Fisher’s exact test). If those patients who had a recurrence were also counted as completers, then completion rates were 73.3% for olanzapine and 67.3% for lithium (p=0.206, Fisher’s exact test).

Patient characteristics are presented in Table 3. Approximately 93% of the patients had a manic index episode, and 26% were experiencing psychotic features. Among patients randomly assigned to a treatment condition, 72.2% were hospitalized for treatment of their index episode at the time they entered into the open-label cotreatment phase (lithium=74.8%; olanzapine=69.6%) (p=0.24, Fisher’s exact test). The most common reasons for discontinuation during the double-blind maintenance period were adverse events, lack of efficacy, and patient decision (Table 2). There were no significant between-group differences in reasons for premature discontinuation. The estimated median time to discontinuation was 303 and 207 days for olanzapine- and lithium-treated patients, respectively, and the time to discontinuation for any reason was significantly earlier for patients receiving lithium ($t^2=5.7$, df=1, p=0.02, log-rank test). Because symptomatic recurrence in a few cases did not necessarily result in immediate discontinuation from the trial, a post hoc tabulation of the number of completers with sustained remission throughout the double-blind phase of the study also shows significantly higher completion with sustained remission for olanzapine (43.3% [N=94 of 217]) than lithium (28.5% [N=61 of 214]) (p=0.002, Fisher’s exact test). If those patients who had a recurrence were also counted as completers, then completion rates were 73.3% for olanzapine and 67.3% for lithium (p=0.206, Fisher’s exact test).

Patient characteristics are presented in Table 3. Approximately 93% of the patients had a manic index episode, and 26% were experiencing psychotic features. Among patients randomly assigned to a treatment condition, 72.2% were hospitalized for treatment of their index episode at the time they entered into the open-label cotreatment phase (lithium=74.8%; olanzapine=69.6%) (p=0.24, Fisher’s exact test). The most common reasons for discontinuation during the double-blind maintenance period were adverse events, lack of efficacy, and patient decision (Table 2). There were no significant between-group differences in reasons for premature discontinuation. The estimated median time to discontinuation was 303 and 207 days for olanzapine- and lithium-treated patients, respectively, and the time to discontinuation for any reason was significantly earlier for patients receiving lithium ($t^2=5.7$, df=1, p=0.02, log-rank test). Because symptomatic recurrence in a few cases did not necessarily result in immediate discontinuation from the trial, a post hoc tabulation of the number of completers with sustained remission throughout the double-blind phase of the study also shows significantly higher completion with sustained remission for olanzapine (43.3% [N=94 of 217]) than lithium (28.5% [N=61 of 214]) (p=0.002, Fisher’s exact test). If those patients who had a recurrence were also counted as completers, then completion rates were 73.3% for olanzapine and 67.3% for lithium (p=0.206, Fisher’s exact test).

**Results**

**Patient Disposition and Characteristics**

A total of 543 patients were enrolled during the open-label period and received lithium/olanzapine cotreatment. Ultimately, 431 (79.4%) achieved symptomatic (protocol-defined) remission criteria during the open-label period and were randomly assigned to double-blind maintenance monotherapy with olanzapine (N=217) or lithium (N=214). The most common reasons for discontinuation during the open-label period were patient decision and adverse events (Table 2). Of the patients achieving symptomatic remission, 91.6% (N=395) achieved syndromic remission. One hundred seventy-one patients completed the double-blind maintenance period, with significantly more olanzapine-treated than lithium-treated patients completing the trial (46.5% [N=101] versus 32.7% [N=70], respectively; p=0.004, Fisher’s exact test). The most common reasons for discontinuation during the double-blind maintenance period were adverse events, lack of efficacy, and patient decision (Table 2). There were no significant between-group differences in reasons for premature discontinuation. The estimated median time to discontinuation was 303 and 207 days for olanzapine- and lithium-treated patients, respectively, and the time to discontinuation for any reason was significantly earlier for patients receiving lithium ($t^2=5.7$, df=1, p=0.02, log-rank test). Because symptomatic recurrence in a few cases did not necessarily result in immediate discontinuation from the trial, a post hoc tabulation of the number of completers with sustained remission throughout the double-blind phase of the study also shows significantly higher completion with sustained remission for olanzapine (43.3% [N=94 of 217]) than lithium (28.5% [N=61 of 214]) (p=0.002, Fisher’s exact test). If those patients who had a recurrence were also counted as completers, then completion rates were 73.3% for olanzapine and 67.3% for lithium (p=0.206, Fisher’s exact test).

**Patient characteristics are presented in Table 3. Approximately 93% of the patients had a manic index episode, and 26% were experiencing psychotic features. Among patients randomly assigned to a treatment condition, 72.2% were hospitalized for treatment of their index episode at the time they entered into the open-label cotreatment phase (lithium=74.8%; olanzapine=69.6%) (p=0.24, Fisher’s exact test). Overall, treatment groups were comparable with respect to demographic and clinical characteristics.**

The mean doses of olanzapine and lithium, respectively, during the open-label period were 13.5 mg/day (SD=4.0) and 1003.3 mg/day (SD=267.0) (mean serum level=0.697 meq/liter, SD=0.14). During the open-label phase, investigators were to titrate the lithium dose to attain therapeutic levels by week 4. Considering the post-titration stabilization period only (from week 4 to random assignment), the mean dose of lithium was 1097.0 mg (SD=277.0), and the mean lithium serum level was 0.76 meq/liter (SD=0.14). For the double-blind period, the mean dose was 11.9 mg

---

**TABLE 2. Treatment Discontinuation Rates Among Bipolar Disorder Patients During Open-Label Acute Cotreatment With Olanzapine and Lithium and Double-Blind Olanzapine or Lithium Monotherapy**

<table>
<thead>
<tr>
<th>Reason for Discontinuation</th>
<th>Open-Label Acute Cotreatment With Olanzapine and Lithium (N=543)</th>
<th>Double-Blind Maintenance Monotherapy (N=431)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Adverse events</td>
<td>34</td>
<td>6.3</td>
</tr>
<tr>
<td>Lack of efficacy</td>
<td>16</td>
<td>2.9</td>
</tr>
<tr>
<td>Patient decision</td>
<td>37</td>
<td>6.8</td>
</tr>
<tr>
<td>Criteria not met/noncompliance</td>
<td>8</td>
<td>1.5</td>
</tr>
<tr>
<td>Sponsor decision</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>Investigator decision</td>
<td>9</td>
<td>1.7</td>
</tr>
<tr>
<td>Lost to follow-up</td>
<td>2</td>
<td>0.9</td>
</tr>
</tbody>
</table>

a Based on the investigator’s choice of a single primary reason for ending participation in the study.

b Number included in analysis of primary outcome.

c Includes patients who chose to discontinue due to perception of satisfactory response.

For four of these patients, the investigator listed the reason as “due to noncompliance.”

---

Incidence of and Time to Mood Episodes

Symptomatic recurrence of any mood episode following remission of mania or depression was observed in 38.8% of lithium-treated and 30.0% of olanzapine-treated patients (Table 5). Statistical noninferiority of olanzapine relative to lithium was established because the 95% confidence interval about the observed 8.8% absolute risk reduction (–0.1% to 17.8%) excludes the predefined margin of noninferiority (–7.3%). Considering pole-specific recurrences, olanzapine and lithium did not differ significantly in the proportion of patients who had a depressive recurrence. However, significantly fewer olanzapine-treated patients had recurrence of manic or mixed episodes compared with lithium-treated patients. Time to symptomatic recurrence to any mood episode was not significantly different between treatments (Figure 1).

Recurrence was further assessed as 1) meeting symptomatic recurrence criteria or hospitalization for a mood episode and 2) meeting DSM-IV criteria for syndromic recurrence after having met syndromic criteria for remission. Rates of recurrence and odds ratios are presented in Table 5. Considering both these criteria, significantly fewer olanzapine-treated patients experienced a mood episode recurrence compared with lithium-treated patients. Furthermore, as shown in Figure 1, time until mood

### TABLE 3. Demographic and Clinical Characteristics of Bipolar Disorder Patients Stabilized With Olanzapine and Lithium Cotreatment Then Randomly Assigned to Double-Blind Olanzapine or Lithium Monotherapy

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Open-Label Acute Cotreatment With Olanzapine and Lithium (N=543)</th>
<th>Double-Blind Maintenance Monotherapy Olanzapine (N=217)</th>
<th>Lithium (N=214)</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>p^a</td>
</tr>
<tr>
<td>Female</td>
<td>290 53.4</td>
<td>113 52.1</td>
<td>115 53.7</td>
<td>0.77</td>
</tr>
<tr>
<td>Caucasian</td>
<td>539 99.3</td>
<td>214 98.6</td>
<td>214 100</td>
<td>0.25</td>
</tr>
<tr>
<td>Manic index episode</td>
<td>503 92.6</td>
<td>202 93.1</td>
<td>202 94.4</td>
<td>0.69</td>
</tr>
<tr>
<td>Psychotic features present</td>
<td>149 27.4</td>
<td>59 27.2</td>
<td>53 24.8</td>
<td>0.58</td>
</tr>
<tr>
<td>History of rapid cycling course</td>
<td>22 4.1</td>
<td>6 2.8</td>
<td>7 3.3</td>
<td>0.96</td>
</tr>
<tr>
<td>Lifetime psychotropic medication use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lithium</td>
<td>401 73.8</td>
<td>161 74.2</td>
<td>160 74.8</td>
<td>0.91</td>
</tr>
<tr>
<td>Antipsychotic</td>
<td>509 93.7</td>
<td>202 93.1</td>
<td>204 95.3</td>
<td>0.41</td>
</tr>
<tr>
<td>Valproate</td>
<td>149 27.4</td>
<td>61 28.1</td>
<td>52 24.3</td>
<td>0.38</td>
</tr>
</tbody>
</table>

#### Number of lifetime mood episodes

<table>
<thead>
<tr>
<th>Mania</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Interquartile Range</th>
<th>p^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0.16</td>
</tr>
</tbody>
</table>

#### Age (years)

| Length of current episode (days)                   | 42.9 | 13.0 | 42.5 | 13.1 | 42.3 | 12.3 | 0.21 | 1.387 | 0.65 |
| Time in remission before randomization (days)      | 37.5 | 37.9 | 37.7 | 39.2 | 37.0 | 33.3 | 0.60 | 1.386 | 0.44 |
| Young Mania Rating Scale total score^d             | 25.8 | 7.2  | 3.6   | 3.5   | 3.9  | 3.8  | 1.8  | 1.387 | 0.18 |
| Hamilton depression scale total score^d            | 5.7  | 4.7  | 1.6   | 2.0   | 1.6  | 1.9  | 1.0  | 1.387 | 0.31 |

#### Incidence of and Time to Mood Episodes

Symptomatic recurrence of any mood episode following remission of mania or depression was observed in 38.8% of lithium-treated and 30.0% of olanzapine-treated patients (Table 5). Statistical noninferiority of olanzapine relative to lithium was established because the 95% confidence interval about the observed 8.8% absolute risk reduction (–0.1% to 17.8%) excludes the predefined margin of noninferiority (–7.3%). Considering pole-specific recurrences, olanzapine and lithium did not differ significantly in the proportion of patients who had a depressive recurrence. However, significantly fewer olanzapine-treated patients had recurrence of manic or mixed episodes compared with lithium-treated patients. Time to symptomatic recurrence to any mood episode was not significantly different between treatments (Figure 1).

Recurrence was further assessed as 1) meeting symptomatic recurrence criteria or hospitalization for a mood episode and 2) meeting DSM-IV criteria for syndromic recurrence after having met syndromic criteria for remission. Rates of recurrence and odds ratios are presented in Table 5. Considering both these criteria, significantly fewer olanzapine-treated patients experienced a mood episode recurrence compared with lithium-treated patients. Furthermore, as shown in Figure 1, time until mood

### TABLE 4. Mean Serum Lithium Levels During Double-Blind Maintenance Monotherapy With Olanzapine or Lithium in Bipolar Disorder Patients Following Stabilization With Olanzapine and Lithium Cotreatment

<table>
<thead>
<tr>
<th>Week</th>
<th>N</th>
<th>Mean Serum Level (meq/liter)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Median</td>
<td>Interquartile Range</td>
</tr>
<tr>
<td>0</td>
<td>207</td>
<td>0.78</td>
<td>0.18</td>
</tr>
<tr>
<td>2</td>
<td>192</td>
<td>0.75</td>
<td>0.18</td>
</tr>
<tr>
<td>4</td>
<td>183</td>
<td>0.77</td>
<td>0.19</td>
</tr>
<tr>
<td>8</td>
<td>175</td>
<td>0.78</td>
<td>0.20</td>
</tr>
<tr>
<td>12</td>
<td>163</td>
<td>0.78</td>
<td>0.22</td>
</tr>
<tr>
<td>16</td>
<td>149</td>
<td>0.77</td>
<td>0.25</td>
</tr>
<tr>
<td>20</td>
<td>136</td>
<td>0.78</td>
<td>0.21</td>
</tr>
<tr>
<td>24</td>
<td>120</td>
<td>0.77</td>
<td>0.25</td>
</tr>
<tr>
<td>28</td>
<td>116</td>
<td>0.76</td>
<td>0.23</td>
</tr>
<tr>
<td>36</td>
<td>101</td>
<td>0.79</td>
<td>0.24</td>
</tr>
<tr>
<td>44</td>
<td>87</td>
<td>0.73</td>
<td>0.27</td>
</tr>
<tr>
<td>52</td>
<td>75</td>
<td>0.79</td>
<td>0.23</td>
</tr>
</tbody>
</table>
episode recurrence was significantly longer for olanzapine-treated patients.

Significantly fewer olanzapine-treated patients (14.3% [N=31 of 217]) were hospitalized for a mood episode during the double-blind period compared with lithium-treated patients (22.9% [N=49 of 214]) (p<0.03, Fisher’s exact test), and time to hospitalization was significantly longer for the olanzapine group (Figure 2). For both groups, the majority of hospitalizations was for recurrence of mania.

Although a documented history of intolerance or lack of response to an adequate trial of olanzapine or lithium was an exclusion criterion, 206 such patients entered the study on lithium regimens. One hundred sixty-four were subsequently randomly assigned to double-blind treatment with lithium (N=80) or olanzapine (N=84). Symptomatic recurrence criteria were met by 46.3% (N=37) of those given lithium and 34.5% (N=29) of those given olanzapine (p=0.152). The impact of lithium use at study entry was examined further to assess the potential bias favoring olanzapine by comparing the differential rates of recurrence among those who were and were not taking lithium at entry. Among patients not taking lithium at study entry, there was a 7.2% recurrence rate advantage for olanzapine (27.1%) over lithium (34.3%). Among patients taking lithium at study entry, there was an 11.8% advantage for olanzapine (34.5%) over lithium (46.3%). The differential advantage was not significantly different (p=0.724, Breslow-Day test). Additionally, the noninferiority of olanzapine relative to lithium can be shown in each lithium-use-at-entry subgroup, since the two-sided 95% confidence intervals around the risk differences did not cover the predefined −7.3% margin of noninferiority.

**Lithium Levels**

Lithium levels were obtained for 211 of 214 patients, and 171 (81%) maintained lithium levels within the therapeutic range or were brought back into the range within the mandated 30-day timeframe. Of the 40 patients whose serum levels were not brought back into the range within the 30-day timeframe, 32 had low serum levels, and eight had high serum levels. Among patients with low serum levels, seven (21.9%) experienced a recurrence, whereas 76 (42.5%) of the 179 within the range or with high serum levels experienced a recurrence.

There was no difference in recurrence rates between lithium-treated patients with high (≥0.8 meq/liter) versus low (<0.8 meq/liter) lithium serum levels (39.0% [N=32 of 82] and 39.5% [N=51 of 129], respectively; p=1.00). Furthermore, in the 83 lithium-treated patients who experienced a recurrence, the mean lithium serum levels were slightly higher (mean=0.78 meq/liter, SD=0.11) than in the 128 lithium-treated patients who did not recur (mean=0.75 meq/liter, SD=0.16), but the difference was not statistically significant (t=1.7, df=209, p=0.09). Considering the 83 patients who met recurrence criteria, the mean lithium level before recurrence was 0.73 meq/liter (SD=0.30); however, for 15 of these (18.1%), the before-recurrence lithium level was ≤0.6 meq/liter. Excluding these 15 patients with low lithium level before recurrence provides an adjusted overall rate of recurrence of 34.2% (68 of 199) for the lithium therapy group compared with 30.0% for the olanzapine group (p=0.40). The two-sided 95% confidence interval on this observed 4.2% risk difference (−4.8% to 13.2%) is consistent with the noninferiority of olanzapine relative to lithium.

We further assessed whether lithium levels were a factor in study completion/disposition, using revised disposition categories in which recurrence superseded any other reported reason for discontinuation. The mean lithium serum level for recurring lithium patients was 0.78 meq/liter (SD=0.11) and was 0.79 meq/liter (SD=0.11) for patients who completed the study in sustained remission. The lowest mean lithium level was found among patients who discontinued due to not meeting protocol criteria or noncompliance (N=10, mean=0.57 meq/liter, SD=0.27). Among other reasonably sized (N>8) disposition groups, mean

---

**TABLE 5. Mood Episode Recurrence Risk During Double-Blind Olanzapine or Lithium Maintenance Monotherapy in Bipolar Disorder Patients Following Stabilization With Olanzapine and Lithium Cotreatment**

<table>
<thead>
<tr>
<th>Recurrence Definition and Mood Episode Type</th>
<th>Olanzapine (N=217)</th>
<th>Lithium (N=214)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Symptomatic recurrence&lt;sup&gt;a&lt;/sup&gt;</td>
<td>65</td>
<td>30.0</td>
</tr>
<tr>
<td>Depression</td>
<td>34</td>
<td>15.7</td>
</tr>
<tr>
<td>Mania</td>
<td>30</td>
<td>13.8</td>
</tr>
<tr>
<td>Mixed</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Symptomatic recurrence&lt;sup&gt;a&lt;/sup&gt; or hospitalization</td>
<td>68</td>
<td>31.3</td>
</tr>
<tr>
<td>Depression</td>
<td>36</td>
<td>16.6</td>
</tr>
<tr>
<td>Mania</td>
<td>30</td>
<td>13.8</td>
</tr>
<tr>
<td>Mixed</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>Syndromic recurrence&lt;sup&gt;b&lt;/sup&gt;</td>
<td>53</td>
<td>26.2</td>
</tr>
<tr>
<td>Depression</td>
<td>28</td>
<td>13.9</td>
</tr>
<tr>
<td>Mania</td>
<td>24</td>
<td>11.9</td>
</tr>
<tr>
<td>Mixed</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

<sup>a</sup> Score ≥15 on Young Mania Rating Scale and/or Hamilton depression scale.  
<sup>b</sup> Random assignment to treatment condition was based on meeting symptomatic remission criteria; not all patients also met the DSM-IV syndromic remission criteria. Consequently, the Ns are smaller for recurrence based on DSM-IV syndromic criteria (for olanzapine, N=202; for lithium, N=193).
lithium levels were 0.72 meq/liter (SD=0.13) for those dis-
continuing due to patient decision (N=25), 0.85 meq/liter
(SD=0.12) for those discontinuing due to an adverse event
(N=11), and 0.75 meq/liter (SD=0.15) for those discontinu-
ing due to lack of efficacy (N=9).

Adverse Events

One patient committed suicide during the open-label
phase of this study. During the double-blind period, two
patients randomly assigned to lithium died. One of these
patients committed suicide, the other died of accidental
causes.

Thirty-four patients (6.3%) discontinued treatment dur-
ing the open-label period due to an adverse event. Com-
mon (≥5%) treatment-emergent adverse events during
this period were increased weight (10.3%), tremor (9.8%),
sedation (7.2%), somnolence (6.8%), and insomnia (5%).
During the double-blind period, adverse events led to the
withdrawal of 41 patients in the olanzapine group (18.9%)
and 55 patients in the lithium group (25.7%). Common or
significant treatment-emergent adverse events occurring
during double-blind monotherapy are reported in Table 6.
We examined the occurrence of depression as a treatment-emergent adverse event in relation to depressive relapse among the olanzapine-treated patients. Of 45 olanzapine-treated patients with treatment-emergent depression, 23 met relapse criteria. Most (73.9%) of these 23 olanzapine-treated patients had Hamilton depression scale total scores ≥15 at the time of the reported emergence of depressive symptoms, and most (65.2%) met relapse criteria within 14 days of the event onset. Similarly, we examined the occurrence of insomnia as a treatment-emergent adverse event in relation to manic symptoms among the lithium-treated patients. Overall, 22/48 patients with insomnia met relapse criteria, but only two cases of relapse were within 1 week of insomnia emergence, and most (63%) had event onset >14 days before relapse. In addition, elevated mood (Young Mania Rating Scale item 4), however, was reduced; 24 patients (51%) had scores of ≥2 (sleeping less than normal by more than 1 hour) near when treatment-emergent insomnia was reported as an adverse event.

**Extrapyramidal Symptom Ratings**

Extrapyramidal symptoms were monitored as patient-reported treatment-emergent events (data not shown), rating scale-defined treatment-emergent events (see Method section), and as mean change in scores on rating scales (data not shown). Irrespective of the means of assessment, changes in and incidences of extrapyramidal symptoms were small and did not differ statistically between treatment groups (Table 7).

**Vital Signs, Weight, and Laboratory Measures**

There were no statistically significant differences between treatments in the incidence rates of potentially clinically relevant changes in vital signs during double-blind therapy. Mean weight gain during the open-label period was 2.74 kg (SD=3.8), and 148 (27.8%) of 532 experienced ≥7% change from baseline. Mean change in weight during the double-blind period was significantly greater for the olanzapine group (mean=1.8 kg, SD=5.8) than in the lithium group (mean=–1.4 kg, SD=5.0) (F=21.2, df=1, 385, p<0.001). Significantly more olanzapine-treated patients had ≥7% increase in weight than lithium-treated patients (29.8% [N=64] versus 9.8% [N=21], respectively) (p<0.001, Fisher's exact test).

No statistically significant differences occurred between treatment groups in the rates of potentially clinically relevant changes in laboratory measures. More detailed analyses on nonfasting glucose and cholesterol outcomes are presented in Table 8. During double-blind treatment, the mean baseline-to-endpoint change in cholesterol was greater for patients treated with olanzapine compared with lithium-treated patients. However, no significant differences occurred in incidence rates of potentially clinically relevant increases in nonfasting glucose or cholesterol between treatment groups.

**Discussion**

This is the first double-blind, randomized, controlled study to investigate the potential of an atypical antipsychotic to prevent recurrence of bipolar disorder in comparison with any active treatment. Olanzapine and lithium did not statistically differ in preventing mood episode recurrence according to symptomatic rating scale criteria. However, olanzapine was significantly more effective than lithium in preventing recurrence of manic and mixed episodes. Olanzapine’s superiority to lithium in the prevention of mania recurrence is important because in a rigorous recently conducted meta-analysis (4), prevention of mania recurrence had been identified as a particular strength of lithium. Prevention of depression recurrence was similar between the treatments. Time until premature discontinuation for any reason occurred significantly ear-
lier for lithium-treated patients; the estimated median time to discontinuation was approximately 100 days sooner than with olanzapine.

Lithium is the most extensively studied mood stabilizer; meta-analyses have indicated that it is superior to placebo in the prevention of relapse (4) and reduces the risk of relapse 3.6-fold (5). Among placebo-controlled trials that have addressed potential bias associated with rapid withdrawal of lithium, rates of relapse with lithium were 31% (Bowden et al.’s 12-month study [6]), 36% (12-month outcomes reported by Prien et al. [13]), and 40.9% (Bowden et al.’s 18-month study [7]). In these studies, lithium was superior to placebo in preventing bipolar disorder relapse among patients who had a manic index episode. Consistent with these reports, our results showing a recurrence rate of 38.8% after 48 weeks of lithium monotherapy further support the prophylactic efficacy of lithium in bipolar disorder.

Three design measures were taken to minimize bias for either treatment. First, patients with a history of nonresponse, or lack of tolerance, to lithium or olanzapine were excluded from the study. Even though a number of patients were taking lithium at study entry, which may have given a potential selection bias advantage in favor of olanzapine in preventing recurrence, the putative advantage difference was not significantly different (p=0.724). Second, the study design included a 4-week taper period as a means of preventing recurrences associated with the abrupt withdrawal of lithium (14). Third, the enriched design of treating patients during their index episode with a combined regimen of lithium and olanzapine ensured that randomly assigned patients were not preselected to respond preferentially to one treatment. Also note that small proportions of mixed-episode bipolar patients or patients with a history of a rapid-cycling course were enrolled in this study, which may reflect the exclusion of patient subtypes who may respond poorly to lithium (15). Last, serum levels of lithium were comparable between those individuals who did and did not have a recurrence; lithium levels also did not appear to decrease at 150 days, a time when recurrence with lithium appeared to accelerate.

Both olanzapine and lithium were generally well tolerated. Few extrapyramidal symptom events, whether measured subjectively or objectively, occurred during this 52-week study. Weight gain was significantly greater in the olanzapine group. The pattern of weight gain was comparable with previous observations (11, 16); most of the weight gain occurred early (open-label period), followed by an additional 1.8 kg gained during the maintenance period.

No significant differences occurred between the groups in mean baseline-to-endpoint changes in nonfasting glucose or incidences of nonfasting glucose levels ≥200 mg/dl. It is noteworthy, however, that this study may not have had sufficient power to determine treatment differences in these adverse events. It is also noteworthy that 43.8% of patients had a nonfasting cholesterol level at screening that exceeded 200 mg/dl. Measurements of the effects of atypical antipsychotics on glucose and cholesterol are important both because of apparent increased rates of diabetes among patients with bipolar disorder and case reports of diabetes among patients treated with atypical antipsychotic agents.

Several limitations in this study warrant discussion. There was no placebo arm in this trial, but both lithium (17) and olanzapine (18) have demonstrated superior relapse prevention relative to placebo. Assessment of previ-
ous response to lithium or olanzapine was collected retrospectively and, as such, may have questionable reliability. Retrospective collection of information is a limitation of all studies that attempt to set exclusion criteria to prevent bias associated with known lack of response to a comparator. This topic is an issue of all contemporary active-controlled clinical trials and deserves more attention by clinical trial methodologists.

In recent meta-analyses examining the effectiveness of lithium for relapse prevention, lithium levels for six trials ranged between 0.5 and 1.4 meq/liter (17). Included in these meta-analyses were two 1973 studies of Prien et al. (13), which reported median lithium levels of 0.7 and 0.8 meq/liter; the 1971 study of Coppen et al. (19), which reported a mean lithium level of 0.93 meq/liter; the 2000 study of Bowden et al. (20), which compared valproate, lithium, and placebo and reported a mean level of lithium of 1.0 meq/liter at day 30; and the 2003 study of Calabrese et al. (21), which compared lamotrigine, placebo, and lithium in patients with an index episode of depression and reported a mean lithium level of 0.8 meq/liter. With the exception of the Bowden et al. study (in which lithium did not separate from placebo), the mean (0.76 meq/liter) and median (0.8 meq/liter) lithium levels reported in this trial are consistent with those reported in the literature. However, it is possible that some patients may not have been maintained on optimum therapeutic levels, which may represent a limitation of the study.

The generalizability of this study is limited to 52 weeks and mainly to individuals with a recent manic episode, since there were few patients with a mixed index episode or a history of rapid cycling, and entry criteria excluded patients with an index episode of depression. In addition, the results can only be generalized to patients stabilized with a combined regimen of olanzapine and lithium and may not be applicable to those stabilized with other mood stabilizers or combinations of mood stabilizers. The interpretation of recurrence on the basis of DSM-IV syndromic criteria may be limited because the definition of recurrence did not include a duration criterion.

It is noteworthy that this study may not have had sufficient power to determine treatment differences in rare adverse events and that assessment of the potential impact of treatment on glucose homeostasis is limited in this study because glucose and lipid measurements were non-fasting. Another limitation is that more than half of the patients withdrew prematurely from the study. High discontinuation rates in maintenance studies are common. Bowden et al. (6) reported dropout rates of 62% and 76% among patients treated with divalproex and lithium for 52 weeks, and dropout rates of 95%, 98%, and 100% were reported for patients treated with lamotrigine, lithium, and placebo for 18 months (7). Finally, the time in remission before random assignment was relatively short; however, when early remitters were excluded from the analyses, risks of recurrence were similar to the previous values.

The results of this trial are consistent with previous studies demonstrating the efficacy of lithium in relapse prevention in bipolar disorder and suggest that olanzapine may also be effective in the prevention of relapse/recurrence in this disorder. Additional independent studies are needed to confirm these results.

Acknowledgments

The following countries and individuals participated in the HGHT clinical trial: Australia: Prof. P. Morris, Dr. R. Newton, Dr. D. Tannenbaum; Austria: Prof. S. Kasper, Dr. M. Schmitz, Prof. C. Simhandl; Belgium: Dr. G. De Bruecker, Dr. H. Bryon, Dr. B. Gillain; Bulgaria: Prof. V. Milanova, Dr. O. Tanchev, Prof. S. Totodorov; Canada: Dr. P. Carr, Dr. L. N. Yaham; Croatia: Dr. V. Jukic, Dr. N. Mandic, Prof. L. Moro; Czech Republic: Dr. E. Bockova, Dr. J. Boucek, Dr. V. Hanuskova, Dr. M. Marsalek, Dr. M. Poricky, Dr. D. Seifertova; Denmark: Dr. N. R. Hansen; Finland: Dr. H. J. Koponen, Dr. I. Larmo, Dr. K. Lehtinen, Dr. V. Nevalainen, Dr. R. Riihikangas; Germany: Dr. A. Berghöfer, A/Prof. P. Bräunig, Prof. E. Haen, Prof. F. Henn, Prof. W. Maier, Prof. A. Marneros, Prof. H. J. Moeller, Prof. M. Schmauß; Hungary: Dr. L. Haraszt, Dr. L. Mod, Dr. G. Ostorharics-Horvath, Prof. Z. Rihmer, Dr. G. Vincze; Ireland: Dr. V. O’Keane, Dr. D. Walshe; Israel: Prof. H. P. Belmaker; Italy: Prof. C. Bellantuono, Prof. F. Bogetto, Prof. G. B. Cassano, Prof. A. Koukopoulos, Prof. C. Maggini, Prof. G. Minnai, Prof. G.M. Muscettola; Lithuania: Dr. B. Burba, Dr. V. Maciuslis, Dr. L. E. Radavicius; the Netherlands: Prof. H. E. Kraan, Dr. J. J. Van Egmond; New Zealand: Prof. T. Silverstone; Norway: Prof. O. Angland, Dr. B. Stubbhaug; Poland: Dr. W. Chrzanowski, Dr. A. Czernikiewicz, Dr. J. Janczewski, Dr. J. Laczkowski, Dr. J.K. Rybakowski; Romania: Prof. P. Boiteanu, Dr. A. I. Grigoriu, Dr. A. S. Ioanes; Russia: Prof. L.M. Bardenstein, Prof. N. G. Neznanov, Prof. G. P. Pantelyevaya, Prof. A. B. Smulevich; Slovenia: Dr. E. Palova, Dr. L. Varuvoska, Dr. L. Vircik; Slovakia: A/Prof. S. Zierhi; South Africa: Dr. C. Grobler; Sweden: Prof. R. Adolfsson, Dr. G. Arnell, Dr. L. Hagstrom, Dr. C. Röllers, Dr. P. Skeppar; Switzerland: Dr. B. Blajev; Turkey: Prof. T. Oral; United Kingdom: Dr. C. M. Bonthala, Dr. A. Gregoire, Dr. D. Patience, Dr. S. Vethanayagam.

Presented in part at the Stanley Foundation Conference on Bipolar Disorder, Freiburg, Germany, Sept. 12–14, 2002; the 41st annual meeting of the American College of Neuropsychopharmacology, San Juan, Puerto Rico, Dec. 8–12, 2002; the 156th annual meeting of the American Psychiatric Association, San Francisco, May 17–22, 2003; and the fifth International Conference for Bipolar Disorders, Pittsburgh, June 12–14, 2003. Received Oct. 30, 2003; revisions received June 17 and Aug. 2, 2004; accepted Aug. 11, 2004. From Lilly Research Laboratories; the Harvard Medical School Department of Psychiatry, McLean Hospital, Belmont, Mass.; the University of Munich Department of Psychiatry, Munich; the Department of Psychiatry, Case Western Reserve University, Cleveland; the Harvard Medical School Department of Psychiatry, Massachusetts General Hospital, Boston; the Department of Psychiatry, University of British Columbia, Vancouver; Former Research Group Clinical Psychopharmacology, Free University Berlin, Germany; Centro Lucio Bini, Rome, Italy; the Department of Psychiatry, University of Pisa, Italy; Mood Disorders Research Unit, Aarhus University Hospital, Aarhus, Denmark; Lilly Area Medical Center, Vienna, Austria; and the Department of Psychiatry, University of Texas Health Sciences Center, San Antonio. Address correspondence and reprint requests to Dr. Tohen, Lilly Research Laboratories, Indianapolis, IN 46285; m.tohen@lilly.com (e-mail).

This study was sponsored by Lilly Research Laboratories.
References

2. Tohen M, Waterman CM, Tsuang MT: Outcome in mania: a 4-year prospective follow-up of 75 patients utilizing survival analysis. Arch Gen Psychiatry 1990; 47:1106–1111
Service Use and Outcomes of First-Admission Patients With Psychotic Disorders in the Suffolk County Mental Health Project

Ramin Mojtabai, M.D., Ph.D., M.P.H.
Daniel Herman, D.S.W., M.S.
Ezra S. Susser, M.D., Dr.P.H.
Nancy Sohler, Ph.D., M.P.H.
Thomas J. Craig, M.D., M.P.H.
Janet Lavelle, M.S.
Evelyn J. Bromet, Ph.D.

Objective: The purpose of the study was to examine the inpatient and outpatient service use and 4-year outcomes of newly admitted psychotic patients during a period of rapid change in the provision of psychiatric services in a well-defined catchment area in New York State in the 1990s.

Method: Subjects were 573 participants of the Suffolk County Mental Health Project. This group comprised patients with psychotic disorders first admitted between September 1989 and August 1995 to 12 inpatient facilities across Suffolk County, N.Y., and followed for up to 48 months. The subjects' service use, course of illness, symptomatic outcomes, suicide risk, homelessness risk, and satisfaction with care were compared across admission years.

Results: The length of inpatient stays decreased significantly across the years. However, the number of outpatient visits and therapy sessions did not vary. Although the patients admitted in later years were more symptomatic at admission to their first hospitalization, their course and outcomes over the follow-up period were not worse and they were not less satisfied with their care, compared with the patients admitted in earlier years.

Conclusions: The clinical characteristics of patients and the role of inpatient care in the management of patients with psychotic disorders gradually changed during the 1990s. These changes, however, were not associated with changes in the use of outpatient services or outcomes. Nevertheless, shorter hospital stays and the presence of more severely ill patients highlight the need for more attention to linkage to aftercare and enhancement of support networks in the community.

(Am J Psychiatry 2005; 162:1291–1298)

The 1990s witnessed many changes in mental health services for patients with severe mental disorders (1). Perhaps the most visible of these changes was the reduced reliance on inpatient treatment. Although this trend began in the early 1960s with the first major wave of deinstitutionalization, it continued well through the 1990s. Between 1988 and 1994, for example, the total number of days of care in mental hospitals declined by 12.5 million days per year (2), a decrease that was only partly offset by an increase of 1.2 million days of psychiatric care in general hospitals. This decline was due mostly to shorter lengths of hospital stays, as the number of psychiatric discharges did not decrease during this period (2).

Policies intended to reduce the length of hospital stays were partly motivated by cost concerns. Many policy makers and clinicians also believed that partial hospitalization and outpatient services would be as effective as inpatient care but would be less restrictive and more conducive to patients’ integration into the community.

Whether and to what extent this shift in locus of care occurred and its effects on clinical and social outcomes of patients with severe mental illness have yet to be fully examined. A study of privately insured patients treated between 1993 and 1995 revealed a paradoxical decrease in the use of outpatient services that accompanied a reduction in inpatient days (3). Such overall reduction in services is a cause for concern, particularly for patients with severe mental disorders. Many of these patients require long-term aftercare, and some require rehospitalization for stabilization or medication adjustment. Thus, reduced use of services may adversely affect the course and outcome of these conditions.

We used data from a longitudinal epidemiological study of first-admission patients with psychotic disorders in the early to mid-1990s in a well-defined catchment area in New York State to examine changes in the mix and volume of services provided to this patient population during this period. We also examined changes in the patients' clinical and social outcomes and in their satisfaction with care. Recruitment that extended over a period of 6 years provided a natural experiment in which systematic variations in usual services and outcomes in this setting could be studied. We focused on first admissions in order to minimize the effect of past treatment history and better reflect the treatment careers of new entrants into the care system.
### TABLE 1. Characteristics of Patients and Index Hospitalizations of 573 Patients Admitted in 1989–1995 to 12 Inpatient Facilities in the Suffolk County Mental Health Project, by Admission Cohort

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>Mean: 30.4</td>
<td>Mean: 29.9</td>
<td>Mean: 29.3</td>
<td>Mean: 29.6</td>
</tr>
<tr>
<td></td>
<td>SD: 9.7</td>
<td>SD: 9.9</td>
<td>SD: 9.9</td>
<td>SD: 8.7</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>48 (46.2%)</td>
<td>46 (45.5%)</td>
<td>34 (46.0%)</td>
<td>23 (32.4%)</td>
</tr>
<tr>
<td>Male</td>
<td>56 (53.9%)</td>
<td>55 (54.5%)</td>
<td>40 (54.1%)</td>
<td>48 (67.6%)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>79 (76.0%)</td>
<td>74 (73.3%)</td>
<td>58 (78.4%)</td>
<td>47 (66.2%)</td>
</tr>
<tr>
<td>Minority</td>
<td>25 (24.0%)</td>
<td>27 (26.7%)</td>
<td>16 (21.6%)</td>
<td>24 (33.8%)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or more</td>
<td>78 (75.0%)</td>
<td>82 (81.2%)</td>
<td>51 (68.9%)</td>
<td>57 (80.3%)</td>
</tr>
<tr>
<td>Did not graduate high school</td>
<td>26 (25.0%)</td>
<td>19 (18.8%)</td>
<td>23 (31.1%)</td>
<td>14 (19.7%)</td>
</tr>
<tr>
<td>Type of insurance at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>33 (32.4%)</td>
<td>43 (43.9%)</td>
<td>24 (33.3%)</td>
<td>17 (24.3%)</td>
</tr>
<tr>
<td>Public (Medicaid, Medicare)</td>
<td>17 (16.7%)</td>
<td>18 (18.4%)</td>
<td>10 (13.9%)</td>
<td>13 (18.6%)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (2.0%)</td>
<td>5 (5.1%)</td>
<td>2 (2.8%)</td>
<td>3 (4.3%)</td>
</tr>
<tr>
<td>None</td>
<td>50 (49.0%)</td>
<td>32 (32.7%)</td>
<td>36 (50.0%)</td>
<td>37 (59.2%)</td>
</tr>
<tr>
<td>Baseline DSM-III-R research diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>30 (28.9%)</td>
<td>26 (25.7%)</td>
<td>26 (35.1%)</td>
<td>22 (31.0%)</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>21 (20.2%)</td>
<td>27 (26.7%)</td>
<td>19 (25.7%)</td>
<td>12 (16.9%)</td>
</tr>
<tr>
<td>Major depressive disorder</td>
<td>15 (14.4%)</td>
<td>16 (15.8%)</td>
<td>14 (18.9%)</td>
<td>8 (11.3%)</td>
</tr>
<tr>
<td>Psychotic disorder</td>
<td>5 (4.8%)</td>
<td>6 (5.9%)</td>
<td>7 (9.5%)</td>
<td>12 (16.9%)</td>
</tr>
<tr>
<td>Other</td>
<td>33 (31.7%)</td>
<td>26 (25.7%)</td>
<td>8 (10.8%)</td>
<td>17 (23.9%)</td>
</tr>
<tr>
<td>Lifetime alcohol/substance disorder</td>
<td>52 (50.0%)</td>
<td>50 (49.5%)</td>
<td>39 (52.7%)</td>
<td>39 (54.9%)</td>
</tr>
<tr>
<td>Type of facility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community</td>
<td>33 (31.7%)</td>
<td>32 (31.7%)</td>
<td>17 (23.0%)</td>
<td>13 (18.3%)</td>
</tr>
<tr>
<td>State</td>
<td>43 (41.4%)</td>
<td>32 (31.7%)</td>
<td>21 (28.4%)</td>
<td>27 (38.0%)</td>
</tr>
<tr>
<td>University hospital</td>
<td>25 (24.0%)</td>
<td>31 (30.7%)</td>
<td>33 (44.6%)</td>
<td>28 (39.4%)</td>
</tr>
<tr>
<td>Veterans’ hospital/other</td>
<td>3 (2.9%)</td>
<td>6 (5.9%)</td>
<td>3 (4.1%)</td>
<td>3 (4.2%)</td>
</tr>
<tr>
<td>Time from onset of disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>to first admission (days)</td>
<td>723.5</td>
<td>1409.8</td>
<td>490.9</td>
<td>1232.7</td>
</tr>
<tr>
<td>Length of hospital stay (days)</td>
<td>41.0</td>
<td>41.0</td>
<td>25.5</td>
<td>41.5</td>
</tr>
<tr>
<td>Clinician-rated psychiatric status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at discharge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In full remission</td>
<td>54 (51.9%)</td>
<td>44 (44.0%)</td>
<td>21 (28.4%)</td>
<td>23 (32.4%)</td>
</tr>
<tr>
<td>In partial remission</td>
<td>21 (20.2%)</td>
<td>19 (19.0%)</td>
<td>39 (52.7%)</td>
<td>37 (52.1%)</td>
</tr>
<tr>
<td>Significant symptoms</td>
<td>1 (1.0%)</td>
<td>5 (5.0%)</td>
<td>5 (6.8%)</td>
<td>3 (4.2%)</td>
</tr>
<tr>
<td>Aftercare referral</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day hospital</td>
<td>3 (2.9%)</td>
<td>5 (5.0%)</td>
<td>3 (4.1%)</td>
<td>1 (1.4%)</td>
</tr>
<tr>
<td>Outpatient clinic</td>
<td>75 (72.1%)</td>
<td>66 (66.0%)</td>
<td>44 (59.5%)</td>
<td>47 (66.2%)</td>
</tr>
<tr>
<td>Private psychiatrist</td>
<td>15 (14.4%)</td>
<td>19 (19.0%)</td>
<td>10 (13.5%)</td>
<td>11 (15.5%)</td>
</tr>
<tr>
<td>Other aftercare</td>
<td>7 (6.7%)</td>
<td>7 (7.0%)</td>
<td>8 (10.8%)</td>
<td>10 (14.1%)</td>
</tr>
<tr>
<td>Loss to follow-up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 6-month interview</td>
<td>4 (3.9%)</td>
<td>4 (4.0%)</td>
<td>3 (4.1%)</td>
<td>8 (11.3%)</td>
</tr>
<tr>
<td>At 24-month interview</td>
<td>9 (8.9%)</td>
<td>6 (6.2%)</td>
<td>7 (10.0%)</td>
<td>9 (13.2%)</td>
</tr>
<tr>
<td>At 48-month interview</td>
<td>12 (12.5%)</td>
<td>15 (16.5%)</td>
<td>9 (13.2%)</td>
<td>11 (16.4%)</td>
</tr>
</tbody>
</table>

a Linear trend across the admission cohorts for categorical variables was assessed with the score test for trend of odds. (For characteristics with more than one category, data for each category were compared to data for all other categories combined.) Linear trend across the admission cohorts for continuous and ordinal variables was assessed with Spearman’s rank-order correlation (r.s).

b Includes schizoaffective disorder and schizophreniform disorder.

c Includes participants who refused the follow-up interview or were not located. The eligible sample for follow-up comprised 552 participants at the 24-month interview and 526 participants at the 48-month interview. Based on a review of all information, including new evidence available at the 6-month consensus diagnostic meeting, 20 participants were deemed not to have been psychotic at baseline and hence were excluded from the 24-month follow-up pool. For the same reason, at the 24-month consensus meeting, an additional 19 participants were excluded from the 48-month follow-up pool. In addition, one participant had died by the 6-month follow-up, and another seven had died by the 24-month follow-up.

*p < 0.05.  **p < 0.01.  ***p < 0.001.
We addressed three specific questions: 1) How did the use of inpatient and outpatient services change among consecutive cohorts of patients admitted in 1989–1995? 2) How did the 48-month course and outcome of these consecutive cohorts of patients change? 3) How did their global functioning and satisfaction with care change?

**Method**

**Subjects**

The data were drawn from the Suffolk County Mental Health Project, a longitudinal epidemiological study of consecutive first admissions to 12 psychiatric facilities in Suffolk County, N.Y., between 1989 and 1995 (4). Briefly, inclusion criteria for the study were age 15–60 years, residence in the county, clinical evidence of psychosis, and both capacity and willingness to provide written informed consent. Exclusion criteria were a psychiatric hospitalization more than 6 months before the current admission, moderate or severe mental retardation, and inability to speak English. Overall, 674 individuals met the inclusion criteria and agreed to participate in the study. We further limited the sample for this study to patients who had no previous hospitalizations (N=600) and who had their first admission between September 1, 1989, and August 31, 1995 (N=573). Although a few patients were admitted to the participating facilities before September 1, 1989, the recruitment sites became fully operational only after this date. Similarly, although the study continued through the early 1996, the recruitment rate in the later months dropped to below 50%, partly because of the extremely short stays of some of the patients.

**Data Collection**

Written informed consent was obtained from the subjects for participation in the study, and their written permission to gather information from medical records and from significant others was obtained. After the baseline interview, the subjects were interviewed by telephone every 3 months and in person at months 6, 24, and 48. Interviews were conducted by trained research interviewers, all of whom were mental health professionals.

**Nonparticipation and Loss to Follow-Up**

The proportion of subjects who agreed to participate and were located for interview at baseline did not vary systematically across the years of the study included in this report (score test for trend=0.10, df=1, p=0.75). Overall, 72.0% (N=674) of the patients referred to the study completed the baseline interview. Patients who were referred to the study but did not agree to participate or were not located were more likely to be older and female and more likely to have their first admission in state or university facilities rather than in community facilities or other types of facilities. We adjusted for these factors in the main analyses.

At the 6-month consensus diagnostic meeting, which included a review of all information available to the project psychiatrists, 20 participants were deemed not to have had a psychotic disorder at baseline and hence were excluded from the 24-month follow-up pool. For the same reason, at the 24-month consensus meeting, an additional 19 participants were excluded from the pool of eligible subjects for the 48-month follow-up. These exclusions had no effect on the results of the study. In addition, one participant had died by the 6-month follow-up, and another seven had died by the 24-month follow-up. Thirty-six (6.3%) of the 573 participants in the follow-up pool at 6 months, 67 (12.1%) of the 552 participants at 24 months, and 94 (17.9%) of the 526 participants at 48 months either could not be located or refused participation and were thus classified as lost to follow-up. The prevalence of loss to follow-up at 24 and 48 months increased across successive...
The type and frequency of outpatient treatment contacts were assessed by using a standardized instrument that was also completed every 6 months.

**Course of Illness, Clinical Outcomes, Suicide, and Homelessness**

The course of illness was rated every 6 months by the interviewers on a modified scale adopted from the Strauss-Carpenter Prognostic Rating Scale (6); the period covered by the rating was the preceding 6 months. Ratings on this scale included “full remission for 3 months or longer,” “full remission for less than 3 months,” “partial remission,” “new episodes during interval,” and “original disorder continued.” For the analyses reported here, these categories were collapsed into the following three categories: full remission (the first two ratings), partial remission, and continuous illness or new episodes (the last two ratings). Full remission was defined as an 8-week period in which the subject was asymptomatic, regardless of treatment status. Partial remission was defined as having some symptoms of the index episode.

Symptoms were rated at baseline and at 6, 24, and 48 months. Ratings were done with the Brief Psychiatric Rating Scale (BPRS) (9), Scale for the Assessment of Positive Symptoms (SAPS) (10), and Scale for the Assessment of Negative Symptoms (SANS) (11).

After each follow-up wave, a search of the National Death Index database was conducted to assess the vital status of participants who were lost to follow-up. On the basis of these data and further information from family members of the deceased participant, suicidal deaths were identified.

Homelessness ratings were conducted at the 24- and 48-month interviews. Ratings were based on self-reports of any homeless nights during the past 2 years.

**Global Functioning**

Global functioning at 24 and 48 months was assessed by using a scale adopted from the SADS-L (8). Ratings included “return to highest level of functioning,” “residual impairment,” “considerable residual impairment,” and “chronic condition or marked deterioration.” Project psychiatrists made these ratings with information from interviews, medical records, interviews with significant others, and rating scales.

**Satisfaction With Care**

The patients’ satisfaction with care was measured at 6 and 24 months by using two questions: 1) How satisfied were you with the quality of services you received? (rated on a scale from 1, quite dissatisfied, to 4, very satisfied) and 2) Did you get the kind of treatment you wanted? (rated on a scale from 1, no, not at all, to 3, yes). Because some patients participated in more than one outpatient program, these questions were asked for as many as three programs. The ratings used here reflect the average across programs. There was no association between the number of programs and the average rating.

**Data Analysis**

Baseline patient characteristics were compared across six cohorts identified on the basis of the date of admission (patients admitted between September 1, 1989, and August 31, 1990, formed the 1989–1990 cohort, those admitted between September 1, 1990, and August 31, 1991, formed the 1990–1991 cohort, and so on). To assess trends across admission cohorts, the score test for trend was used for categorical variables and Spearman's rank-order correlation was used for ordinal variables.

Patterns of treatment, illness course, and symptomatic outcomes over the 48 months were compared across admission cohorts by using generalized estimating equations (12). All generalized estimating equation analyses were adjusted for age, gender, race, education, baseline research diagnosis, and facility type at first admission. Interaction terms for admission cohort with fol-
low-up time were also entered into the models. In addition, analyses of service use were adjusted for updated insurance type (insurance type was assessed at baseline and at 6 and 24 months). In models that showed a statistically significant linear relationship between admission cohort and outcome, we further searched for a possible nonlinear relationship by testing for a quadratic term for admission cohort. Generalized estimating equation analyses were conducted with the Stata 7 xtgee routine (13).

Results

Patient and Index Treatment Characteristics

The patients recruited in the six admission cohorts were similar on most sociodemographic and clinical characteristics (Table 1). The proportion of patients recruited from state facilities declined across the admission cohorts, mainly because of downsizing of the major adult state hospital in the county. The length of first hospital stays declined across admission cohorts. Although the reduction in length of stay occurred in all facility types, it was particularly dramatic for state facilities (Figure 1).

Score test for trend showed a statistically significant trend in the proportion of patients discharged “in partial remission” or “having significant symptoms” across successive cohorts (Table 1). The trend for patients discharged “in partial remission” was largely due to a dramatic increase in such ratings between 1991–1992 and 1992–1993. In subsequent cohorts, however, the proportion of such ratings gradually declined. There was also a systematic increase in the proportion of patients discharged to day hospitals (Table 1).

Course of Treatment

Overall, 43% of the patients were rehospitalized at least once during the 48 months of follow-up (the median number of rehospitalizations among those rehospitalized was two, with a range from one to 12). In the generalized estimating equation analyses, the number of inpatient days over the 48 months declined across admission cohorts (B=-1.84, SE=0.49, z=3.78, p<0.001). However, the generalized estimating equation analyses for the number of rehospitalizations assessed every 6 months revealed no significant variations across cohorts. Thus, the decline in the number of inpatient days was likely due to reduced lengths of stay, not reduced frequency of hospitalizations.

The reduction in inpatient days was not associated with increased use of outpatient services. In the generalized estimating equation analyses, the number of day treatment, individual therapy, medication, and overall outpatient visits did not systematically vary across admission cohorts. Furthermore, generalized estimating equation analyses of the global ratings of course of outpatient treatment (conducted every 6 months) suggested a decline across admission cohorts in the proportion of patients in treatment who received “continuous treatment” (adjusted odds ratio=0.90, 95% confidence interval [CI]=0.82–0.99, z=2.23, p<0.03) and an increase in the proportion of patients who received “several brief periods of treatment” (adjusted odds ratio=1.25, 95% CI=1.05–1.49, z=2.50, p<0.02). The proportion of patients with “consultation/brief periods” of outpatient treatment did not change systematically across admission cohorts nor did the proportion of those who received any outpatient treatment versus none.

Course of Illness, Clinical Outcomes, Suicide, and Homelessness

Illness course and clinical outcomes for the most part did not vary systematically across admission cohorts. The generalized estimating equation analyses revealed an increase in the proportion of patients rated as being in “full remission” across admission cohorts (adjusted odds ratio=1.11, 95% CI=1.01–1.21, z=2.17, p=0.03). Further logistic regression analyses revealed that the difference across cohorts was limited to the 6-month assessment (adjusted odds ratio=1.13, 95% CI=1.02–1.26, z=2.28, p<0.03), and there were no significant differences across cohorts at the later assessment points. There were also no systematic variations in the proportion of patients with ratings of “partial remission” or “new episodes during interval/original disorder continued.”

The comparison of symptom severity in the generalized estimating equation analyses revealed no significant differences in the BPRS and SANS scores. However, analysis of the SAPS scores revealed a nonsignificant trend for higher levels of positive symptoms in later cohorts (B=0.03, SE=0.02, z=1.89, p=0.06). Further analyses using linear regressions revealed that the difference in SAPS scores across admission cohorts was limited to the baseline assessment (B=0.06, SE=0.02, z=3.36, p=0.001), and there were no systematic variations at follow-up assessments.

Five (0.9%) of the 573 participants committed suicide over the 48 months of the study (two in the 1990–1991 cohort and one each in the 1989–1990, 1991–1992, and 1993–1994 cohorts). There was no systematic trend in this variable across admission cohorts (score test for trend=1.44, df=1, p=0.23).

At 48 months, 77 (17.1%) of the 450 participants with follow-up data on housing status reported episodes of homelessness since the first discharge. There was no systematic trend in self-reported homelessness across admission cohorts (score test for trend=0.61, df=1, p=0.44).

Global Functioning

Many patients were rated as having “returned to highest level of functioning” at follow-up (44.8% at 24 months and 44.1% at 24 months). In the generalized estimating equation analyses, a significantly higher proportion of patients in the later admission cohorts had this rating (adjusted odds ratio=1.35, 95% CI=1.13–1.63, z=3.20, p<0.001) and a smaller proportion of patients had a rating of “residual impairment” or “considerable residual impairment” (adjusted odds ratio=0.79, 95% CI=0.65–0.96, z=2.33, p=0.02). The proportion of patients with a rating of “chronic condi-
tion or marked deterioration” did not vary systematically across cohorts.

**Satisfaction With Care**

No differences across admission cohorts were found for satisfaction with care. Overall, 22.6% of the patients at 6 months and 43.7% at 24 months stated that they were “very satisfied” with their services. Also, 40.5% of the patients at 6 months and 64.8% at 24 months stated that they had received the kind of treatment that they wanted. The increase in the proportion of satisfied patients with time may be an artifact of treatment dropout of unsatisfied patients, as only the patients who received treatment during the interval were asked about satisfaction with care. Further analyses showed that a low level of satisfaction at 6 months was associated with a higher likelihood of dropping out of treatment during the next 6 months (satisfaction with services: odds ratio=1.65, 95% CI=1.21–2.24, z=3.18, p=0.001; receiving the kind of treatment that was wanted: odds ratio=2.66, 95% CI=1.78–3.98, z=4.76, p<0.001). Similarly, a low level of satisfaction at 24 months was associated with dropping out of treatment during the next 6 months (satisfaction with services: odds ratio=2.24, 95% CI=1.36–3.68, z=3.16, p=0.002; receiving the kind of treatment that was wanted: odds ratio=3.60, 95% CI=1.95–6.65, z=4.09, p<0.001).

**Discussion**

During the years of study, the mental health care system in Suffolk County, N.Y., and across the United States underwent drastic changes. One major element of these changes was the reduction in the length of inpatient stays. In New York State, these changes were expedited by a new policy initiative—the Community Reinvestment Act of 1993—that was intended to divert funds from inpatient care to outpatient and community-based services (14).

Changes in the use of inpatient services are reflected in our data. The average length of inpatient stays declined drastically across admission cohorts spanning the 1989–1995 period. We also observed a systematic change in the characteristics of first-admission patients hospitalized in Suffolk County across the 6 admission years. Patients in later cohorts had more severe positive symptoms at admission. This difference was probably due to changes in admission policies over this period, as patients who were less severely ill were increasingly less likely to be admitted into hospitals (2).

Both the reduction in the length of hospital stays and the change in the characteristics of patients suggest a shift in the role of inpatient care in the management of severely mentally ill patients. Inpatient services were increasingly used for short-term emergency management of more severely ill patients, and patients with less severe illness were shifted to less intensive settings. Probably as a result of these changes, patients discharged in later years may have been more symptomatic. This pattern was not very clear in our data, as the trend for patients discharged “in partial remission” was largely due to a dramatic increase in such ratings between 1991–1992 and 1992–1993, and, in subsequent cohorts, the proportion of such ratings gradually declined. However, this trend was clearly shown in another study of three cohorts of depressed inpatients discharged between years 1988 and 1996, in which patients who were admitted in later years and who had shorter stays had more symptoms and a lower level of functioning after discharge (15).

Despite the reduced length of inpatient stays over the period of the study, no corresponding increase in the use of outpatient services was found. Although more patients in later admission cohorts were referred to day treatment, this pattern did not translate into increased use of day treatment services. Moreover, continuity of outpatient care, rated globally, did not improve in later admission cohorts. If anything, fewer patients in later cohorts received continuous outpatient treatment. These findings are consistent with the results of a study of a national cohort of privately insured patients that also found no increase in utilization of outpatient services after reduction in inpatient service use (3).

It is noteworthy, however, that the patients admitted during later years experienced a speedy symptomatic recovery after discharge from their first admission and by the 6-month follow-up had symptom measures that were virtually indistinguishable from those of the patients admitted in earlier years. There was also no evidence that the course of illness in later admission cohorts was poorer than that in earlier cohorts. In fact, global measures showed a puzzling trend for patients admitted in later years to function better than those admitted in earlier years.

Although we do not have a ready explanation for these findings, it seems plausible that changes in the structure and content of services in the early to mid-1990s and, most importantly, the drastic reduction in the length of inpatient stays in this period did not adversely affect patient outcomes in the short run. We also did not observe any meaningful trends in rates of suicide and homelessness across admission cohorts. Finally, among patients who remained in care, satisfaction with services did not vary systematically across cohorts.

During the course of the study, admissions to state facilities and, as a result, the proportion of state facility patients in this sample declined dramatically. Because patients admitted to state facilities traditionally have fewer resources and experience poorer course and outcomes, the smaller numbers of such patients in later cohorts could potentially confound the results. However, when the analyses were conducted separately for participants recruited from state facilities and from other facilities, the results were essentially similar to the main results reported here.

As the total number of inpatient psychiatric admissions in Suffolk County declined during the years of the study
(16), the proportion of patients with early psychotic disorders who were admitted also likely declined in later years. Many of these patients may have received care in less intensive settings or in criminal justice settings. However, it is unlikely that sample selection could explain the findings of the study because the patients admitted in later years, if anything, appeared to be more severely ill than those admitted in earlier years.

Another possible explanation for the findings of similar course and outcome at follow-up across admission cohorts despite the more severe presentation at baseline and shorter stays in later cohorts is that the potentially negative effects of these factors were offset by the possible improvements in the content of outpatient services in later years, including the introduction of atypical antipsychotic medications. In our sample, only 19.6% of the patients ever received such medications over the 48 months. Repeating the analyses after excluding these participants produced results similar to those reported here; thus, the findings cannot be attributed to the use of these medications. Nevertheless, changes in other aspects of outpatient care remain a possible explanation to be explored in future research. For instance, some evidence suggests a shift toward more time-limited and behavioral psychosocial interventions during this period at least in community mental health centers (17).

It is also plausible that the reduced length of inpatient stays had a positive effect on the course and outcome of psychotic disorders. Past research on the relationship between length of inpatient stays and clinical and social outcomes produced conflicting results (15, 18–22). Perhaps most relevant to the present study are the results of the McLean First-Episode Psychosis Project, which recruited patients in a time frame similar to the Suffolk County Mental Health Project. That study also recorded a dramatic reduction in average length of stay during the study period (22). But neither time to syndromal recovery nor the proportion of patients attaining syndromal recovery by 2 years varied systematically across the admission cohorts.

In interpreting the results of our study, some limitations should be considered. First, although the Suffolk County Mental Health Project obtained consensus longitudinal diagnoses, we used only the baseline diagnoses in this report to limit the potential effect of course of illness on diagnostic decisions. It is noteworthy, however, that adjustment of the analyses for the 24-month consensus diagnoses did not substantially change the results of the study. Furthermore, the results of analyses after stratification of the data based on the 24-month diagnoses (schizophrenia versus other) were mainly consistent with the results reported here. One notable difference in the subgroup of participants with a 24-month diagnosis of schizophrenia was higher SANS scores in later cohorts. However, this variation was limited to the baseline assessment and did not persist at later assessment points, which suggests that the participants with schizophrenia admitted in later years were more symptomatic at baseline. Second, the scope of measures of service use in this study was limited. Future studies need to examine other domains, including the process and the quality of care (17), use of informal care providers, and use of other services in the community. Finally, future studies also need to examine any possible shift of the burden of care to the criminal justice system in this time frame.

In conclusion, the results of this study contain a mixed message for clinicians and policy makers. On the one hand, the shorter hospital stays and higher likelihood of partial remission or nonremission at the time of discharge call for more attention to provision of community-based support services, including reliable linkage mechanisms that enhance continuation of aftercare in outpatient settings. Simple linking interventions (23) and focused case management programs (24) have shown promising results. The prominent role of families in the care of patients with severe mental disorders also calls for greater attention to supportive and educational family interventions.

On the other hand, it is reassuring to know that shorter hospital stays did not negatively affect the short-term course and outcome of psychotic disorders and that for most patients in the early course of illness, resources available in the community provided adequate substitutes for hospital care.

Received April 29, 2003; revisions received Dec. 10, 2003, and March 31, 2004; accepted Aug. 9, 2004. From the Department of Psychiatry, Beth Israel Medical Center; the Departments of Psychiatry and Epidemiology, Mailman School of Public Health, Columbia University, New York, New York; Montefiore Medical Center and Albert Einstein College of Medicine, Bronx, N.Y.; Department of Veterans Affairs, Office of Quality and Performance, Washington, D.C.; Hillside Hospital, Glen Oaks, N.Y., and the Department of Psychiatry and Behavioral Science, State University of New York at Stony Brook, Stony Brook, N.Y. Address correspondence and reprint requests to Dr. MojiTabai, Department of Psychiatry, Beth Israel Medical Center, First Ave. at 16th St., New York, NY 10010; rm322@columbia.edu (e-mail). Supported in part by NIMH grants MH-01754 and MH-44801.

References

10. Andreasen NC: Scale for the Assessment of Positive Symptoms (SAPS). Iowa City, University of Iowa, 1984
11. Andreasen NC: Scale for the Assessment of Negative Symptoms (SANS). Iowa City, University of Iowa, 1983
Approximately 1.25 million individuals received burn injuries in the United States in 1992 (1). Fire is the third leading cause of unintentional injuries in children (2). Not surprisingly, burns were among the first types of trauma identified as leading to traumatic stress reactions in adults, and early burn research contributed to the original classification of posttraumatic stress disorder (PTSD) in DSM-III (3–5). Early studies of burned children alerted researchers and clinicians to the importance of attending to their psychiatric needs (6, 7). In one of the first studies of the prevalence of psychiatric diagnoses, 30% of the children interviewed more than 6 months after a severe burn had met the DSM-III criteria for PTSD at some point after their burn. These authors also reported higher prevalences of overanxious disorder, phobias, and enuresis in the children with burns than in a nonburn comparison group (8, 9). However, as little is known about the etiology of PTSD in children with burns, there is a great need for risk factor studies. Further, as burned children are accessible for study shortly after their trauma and, by virtue of their ongoing need for medical and surgical care, can be readily followed over time, they are ideally suited to prospective, longitudinal research designs. Such designs are necessary for the advancement of the understanding of risk factors for PTSD (10). In this study we attempted to build a model of risk factors for PTSD symptoms in burned children by assessing them shortly after the burn and then 3 months following this assessment.

Research has focused on the anxiety/arousal and the dissociative symptoms expressed in the acute aftermath of a trauma. The importance of these symptom clusters is formalized in DSM-IV with the addition of acute stress disorder. Acute stress disorder is the psychopathological response in the immediate aftermath of a traumatic event that occurs until 1 month following the trauma and includes both anxiety and dissociative symptoms as necessary components of this diagnosis. Daviss and colleagues (11) and Koplin Winston and colleagues (12) reported a broad range of acute stress disorder symptoms in separate cohorts of injured children. Numerous studies have documented the higher prevalence of PTSD in those initially diagnosed with acute stress disorder (13–15). There is, however, considerable controversy over which symptom cluster, anxiety or dissociation, is more predictive of PTSD and, regarding the diagnosis of acute stress disorder, whether dissociation should be included as a distinct symptom cluster at all. If dissociative symptoms are required in the diagnosis of acute stress disorder, many individuals who suffer from only anxiety symptoms may not receive adequate attention (16, 17). Further, it has been argued that dissociative symptoms are not necessarily more predictive of PTSD than are anxiety symptoms (16). On the other hand, evidence suggests that individuals who dissociate around the time of trauma are at high risk of developing PTSD (18–21). The controversy over the relative importance of anxiety and dissociative symptoms is also important for another reason: these two groupings of symptoms may be the phenotypes of different biobehavioral systems related to PTSD (22, 23). Clarity regarding the relationship between these groups of symptoms and their biological underpinnings may be critical for identifying children at risk and refining immediate interventions following a trauma.
PTSD IN CHILDREN WITH BURNS

TABLE 1. Pretrauma, Trauma, Peritrauma, and Posttrauma Variables for 72 Children With an Acute Burn

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretrauma: age (years)</td>
<td>11.20</td>
<td>3.51</td>
</tr>
<tr>
<td>Trauma: total body surface area burned (%)</td>
<td>17.58</td>
<td>17.92</td>
</tr>
</tbody>
</table>

- Separation anxiety (score on Multidimensional Anxiety Scale for Children [25], possible range=0–27) | 8.21 | 5.25 |
- Dissociation (score on Child Stress Disorders checklist [27], possible range=0–16) | 3.19 | 2.90 |
- Pain (score on Colored Analogue Pain Scale [26], possible range=0–10) | 2.12 | 2.24 |
- Posttrauma: PTSD symptoms 3 months after burn (score on Child PTSD Reaction Index [24], possible range=0–80) | 16.82 | 13.13 |

Of the anxiety symptoms, we regard separation anxiety as particularly important for the population of children hospitalized with burns. The nature of burn injury and its long and stressful hospital course frequently mean separations of the child and the parents at a time when the child has a great need for their help and comfort.

In the current study we used path analytic techniques to evaluate the ways in which pretrauma variables, trauma characteristics, and reactions in the immediate aftermath of a trauma are related to later PTSD symptoms. As described, anxiety and dissociative reactions in the wake of a trauma have frequently been identified as important components of the acute trauma response, although there is considerable controversy over which is more strongly related to deleterious long-term outcomes. Consequently, a main goal of the current study was to assess the relative importance of anxiety and dissociation in the immediate aftermath of a burn.

Method

Participants

The participants were drawn from a group of children admitted to Shriners Burns Hospital in Boston for an acute burn. All children ages 7 to 17 years were eligible to participate unless they or their parents did not speak sufficient English to complete the study instruments. Of 116 eligible children, 72 (62%) participated. Of the 44 children who did not, 26 had families that declined to participate and 18 were discharged before we were able to obtain consent. The mean age of the participants was 11.20 years (SD=3.51); 24 were girls and 48 were boys. The average length of stay was 25 days (SD=23). The mean amount of body surface area burned was 17.58% (range=1%–85%). The children were interviewed an average of 10 days after admission (range=2–26 days).

Procedures

Within 3 days of hospitalization or when the child was considered medically stable (e.g., did not have a delirium, did not have an active infection, and was not receiving mechanical ventilation), the child and his or her parents were approached by one of the investigators (C.L.) and introduced to the study. After complete description of the study, written informed consent and assent were obtained from the parents and child, respectively. On the same day that consent and assent were obtained, the child was interviewed and answered questions about his or her traumatic stress responses and the child’s primary nurse answered questions about the child’s dissociative symptoms (interview and questionnaires are described in the following).

Follow-up assessments consisting of the same interviews and questionnaires were completed by a trained research associate (N.C. or E.H.) at the participant’s home 3 months later.

Measures

Child measures. The Child PTSD Reaction Index (24) is a 20-item semistructured interview that assesses posttraumatic symptoms in children. Its interrater reliability is high (Cohen’s kappa=0.88). Its validity is supported by the finding that children who are known to have PTSD have much higher scores on this instrument (24). The score on the Child PTSD Reaction Index was the main dependent variable in this study.

The Multidimensional Anxiety Scale for Children (25) is a 39-item self-report measure of pediatric anxiety symptoms. In a psychometric study of the scale, the mean intraclass correlation coefficients at 3 weeks and 3 months were 0.79 and 0.93, respectively, demonstrating satisfactory to excellent test-retest reliability (25). Factor analytic studies of the scale have shown a variety of independent scales, including “harm avoidance,” “separation anxiety,” and “physical symptoms.” The separation anxiety scale of the Multidimensional Anxiety Scale for Children was our index of separation anxiety.

The Colored Analogue Pain Scale (26) is a pocket-sized visual analogue instrument on which the child slides a marker along a 10-cm line that shows an increasing intensity of red color corresponding to increased intensity of current pain. This instrument has been used with many groups of children who have pain. Children with more painful syndromes score higher on this instrument than do children with less painful syndromes. The Colored Analogue Pain Scale has been found to be easier to administer than other visual analogue scales.

Nurse measures. The numbing and dissociation scale of the Child Stress Disorders Checklist (27) is a measure of acute dissociative symptoms based on observer report. This 8-item scale assesses the dissociative dimension of the child’s acute and posttraumatic stress response. Internal consistency was found to be 0.75 (Cronbach’s alpha). Test-retest reliability, calculated by correlating scores reported by the parents 2 days apart, was found to be 0.72 (intraclass correlation). In the current study, the child’s primary nurse completed the Child Stress Disorders Checklist in regard to the child’s dissociative symptoms.

The total body surface area burned was the percentage recorded by the attending surgeon in the child’s medical record.

Data Analysis

We used a path analytic strategy similar to that used by Shalev et al. (20) in another prospective study of acutely traumatized individuals. As this strategy is based on a prospective longitudinal method, the directionality of many of the paths was constrained by the time at which the variables were assessed. Accordingly, we divided variables into the following: 1) PTSD symptoms (our main dependent variable, derived from the Child PTSD Reaction Index), 2) posttraumatic variables (variables assessed at the 3-month follow-up), 3) peritraumatic variables (variables assessed shortly after the trauma), 4) trauma exposure variable (percentage of body surface area burned), and 5) pretrauma variables (variables related to the child or family from before the trauma).

A series of hierarchically nested ordinary least squares multiple regression analyses were used to estimate direct and indirect effects among variables. The first step was to predict the dependent variable (PTSD symptoms). From the remaining variables, we chose combinations of variables that accounted for a higher percentage of the variance in PTSD symptoms (high $R^2$ value), guided by our theoretical model of PTSD and constrained by the strength of bivariate relationships (all bivariate relationships with

http://ajp.psychiatryonline.org
r>0.30 and p<0.01). Accordingly, we chose two primary variables (i.e., separation anxiety and dissociative symptoms measured in the hospital) that together accounted for 59% of the variance of PTSD (R^2=0.59). Once these two variables were chosen, antecedent variables could be identified in order to place separation anxiety and dissociation in the roles of mediators. Again, we were guided by theory and the strength of the bivariate relationships, and we were further constrained by temporal relationships. The direction of the relationships chosen had to make temporal sense (e.g., peritraumatic dissociation could not lead to the size of the burn). In this way we generated a network of associations among variables that, as we will discuss, adds important understanding to the unfolding of PTSD over time.

Problems related to missing data can reduce the number of subjects for a particular analysis to a less than optimal level. We employed the statistical package M Plus 2.1 (28) with full information maximum likelihood estimation to retain the full number of subjects for each analysis. M Plus is preferred since it is able to use the full information maximum likelihood procedure in concert with the Satorra-Bentler correction for nonnormal data. (See the work of McArdle [29] and Graham et al. [30] for discussions of the advantages of maximum-likelihood-based methods for incomplete data over more traditional listwise and pairwise deletion procedures.)

Results

Table 1 presents the mean values and standard deviations for the variables used in our path analysis, and Table 2 presents a correlation matrix of these variables.

Figure 1 illustrates the network of associations among variables. The results of the path analysis indicate two direct pathways to PTSD. These pathways are from 1) acute separation anxiety (beta=0.68) and 2) acute dissociation (beta=0.26). In addition, separation anxiety served as a mediator between the age of the child and PTSD and between pain measured shortly after the burn and PTSD. Moreover, both separation anxiety and dissociation mediated the relationship between total burn area and PTSD. Together these pathways account for 59% of the variance in PTSD. The overall model yielded strong fit indices (χ^2=2.0, df=6, p=0.92; comparative fit index=1.00; Tucker-Lewis index=1.00; and root mean square error of approximation=0.00).

Discussion

Two pathways to PTSD symptoms were found. One pathway was mediated by separation anxiety, the other by acute dissociative responses. The magnitude of the trauma, measured by the size of the burn, was not related to PTSD directly but exerted its influence indirectly through both pathways. The pathway mediated by separation anxiety was influenced by the acute pain response. This pathway was also influenced by the size of the burn and was inversely related to the age of the child. The pathway mediated by the acute dissociative response was influenced only by the size of the burn.

The independence of the anxiety and dissociation pathways to PTSD suggests the possibility that different biobehavioral systems contribute to PTSD. A number of researchers have drawn connections between the arousal/anxiety symptoms and the sympathetically mediated fight-or-flight response and between the symptoms of dissociation and the parasympathetically mediated “freeze” or “immobilization” response (22, 23, 31, 32). For example, Perry and colleagues have described the hierarchical response to threat in biological theories of children’s responses to trauma (22, 23). They described the initial fight-or-flight hyperarousal response in a child faced with an immediate threat and the following freeze-or-surrender immobilized response when the child cannot diminish the threat by means of the fight-or-flight response. The authors suggested that this freeze-or-surrender response
occurs when a child is confronted by extreme threat and is helpless to respond. This response is phenotypically observed as dissociative symptoms (22). The fight-or-flight response is controlled by the sympathetic/HPA axis system, and the freeze/immobilization response is controlled by the parasympathetic nervous system (33). Porges has described the evolutionary foundations of this sequentially overwhelmed autonomic nervous system (33). Further, Bowlby (34, 35) has noted the critical interpersonal components of this threat response system in humans and nonhuman primates.

**Anxiety/Arousal Pathway to PTSD**

Bowlby stated, "Of the many fear arousing situations that a child, or older person, can foresee, none is likely to be more frightening than the possibility that an attachment figure will be absent or...unavailable when needed" (35, p. 201). In this study, the child’s level of acute separation anxiety was directly related to PTSD symptoms. Separation anxiety was influenced by the child’s age, the size of the burn, and the degree of pain experienced. As described, the nature of burn injury and the prolonged recovery process during hospitalization mandate stressful separations between the child and parents at a time when the child has a great need for their help and comfort. Thus, the children who experienced the most anxiety on separations were more likely to develop posttraumatic symptoms. Separation anxiety in children may be related to the evolutionarily driven reaction to the anticipated loss of the mother and of her protective function (36). As survival is a frequent concern of children on burn units, the need for soothing and reassurance from a parent is felt intensely by most burned children. This extreme arousal response to being burned, in pain, and alone is consistent with our data.

It is noteworthy that burn trauma has been considered a relatively impersonal trauma, as it is often not caused by another person. The importance of separation anxiety in the development of PTSD suggests that burn trauma has a significant interpersonal component. Specifically, burn trauma and the emotional distress underlying the question “Who will help me?” are integrally connected. This evidence thereby supports the notion that all trauma has an interpersonal element.

**Dissociative Pathway to PTSD**

Krystal has stated, “The switch from anxiety to the cataleptoid response is the subjective evaluation of impending danger as one that cannot be avoided or modified. With the perception of fatal helplessness in the face of destructive danger, one surrenders to it” (37, pp. 114–115). In this study, the degree of dissociative symptoms measured shortly after the burn was found to be a direct predictor of PTSD symptoms. This replicates the results of many studies (18, 19), but we believe that it is the first to document this effect in children with burns. The independence of dissociation from separation anxiety is consistent with the hypotheses of Perry and others that the anxiety/arousal component of PTSD may be a phenotype of the sympathetically mediated fight-or-flight response, whereas the dissociative symptoms may be a phenotype of the parasympathetically mediated immobilization or freezing response. This response is described as phylogenetically very old and is characterized by the lack of vagal tone, bradycardia, and shutting down of responses, in order to conserve resources to maximize the chance of survival during situations of extreme threat. This type of response is thought to occur after exhaustion of other defensive behaviors, such as the fight-or-flight response, which is mediated by the sympathetic nervous system (33, 38, 39). Changes in vagal tone, a well-accepted marker of parasympathetic activity, have been associated with PTSD (40–42). It may be that situations of extreme threat lead to the parasympathetically mediated shutting down of emotional responses, phenotypically observed as dissociative symptoms and prospectively related to PTSD. Parasympathetic nervous system activation was not, however, directly measured in this study.

It is notable that no direct relationship between separation anxiety and dissociation was found in our model. In fact, the bivariate association (Table 2) was negligible (r=0.13). As dissociative responses are hypothesized to follow anxiety/arousal responses when the fight-or-flight approach is ineffective, it would follow that anxiety would lead to dissociation. This dependence of dissociation on anxiety/arousal, but not vice versa, is reported in the accompanying article in this issue of the *Journal* (43). It is possible that variables that were not assessed (such as psychophysiology) may be mediating this relationship.

Regarding the controversy of whether acute anxiety or dissociation is the more important predictor of PTSD, our data suggest that both anxiety and dissociative symptoms independently contribute to the risk for PTSD.

** Limitations **

This study is limited by a relatively small number of subjects and a short longitudinal follow-up. The findings of two independent pathways to PTSD, and our interpretation that they suggest discrete biobehavioral systems, can be illuminated by studies that include psychophysiological and neuroendocrine measures.

** Clinical Implications **

These data have important implications for PTSD treatment and prevention. If symptoms of separation anxiety, pain, and dissociation can be identified in the acute aftermath of a trauma and strongly contribute to the risk of PTSD, then it is critical to assess these symptoms and to intervene accordingly. There are well-described psychosocial and biological interventions for anxiety, pain, and dissociation. Given the particular importance of separation anxiety, burn hospitals (and all pediatric intensive care units) must make every effort to keep parents and children...
together and to work with parents to increase their capacity to comfort their children. It is noteworthy that opiates are among the strongest inhibitors of the distress cry in young animals upon separation from mothers (35, 44). We have previously reported that the dose of morphine received by burned children in the hospital diminishes PTSD symptoms over time (45). It is possible that one mechanism for the diminution of PTSD over time in these children is the effect of the morphine on separation anxiety, as well as pain. If so, it would be a very specific intervention for one of our two pathways to PTSD.

Presented at the 18th annual meeting of the International Society for Traumatic Stress Studies, Baltimore, Nov. 7–10, 2002. Received Dec. 12, 2003; revision received July 20, 2004; accepted Aug. 2, 2004. From the Department of Child and Adolescent Psychiatry, Boston University School of Medicine; the Department of Psychiatry, Shriners Burns Hospital, Boston; and the Department of Psychiatry, Bronx VA Medical Center, Bronx, N.Y. Address correspondence and reprint requests to Dr. Saxe, Department of Child and Adolescent Psychiatry, Boston University School of Medicine, Dowling 1 North, 1 Boston Medical Center Place, Boston, MA 02118; glenn.saxe@ bmc.org (e-mail).

Supported by NIMH grant R01 MH-57370 and by Substance Abuse and Mental Health Services Administration grant U79 SM-54305 (Dr. Saxe).

References

Pathways to PTSD, Part II: Sexually Abused Children

Julie B. Kaplow, Ph.D.
Kenneth A. Dodge, Ph.D.
Lisa Amaya-Jackson, M.D., M.P.H.
Glenn N. Saxe, M.D., F.R.C.P.

Objective: The goal of this research was to develop and test a prospective model of posttraumatic stress symptoms in sexually abused children that includes pretrauma, trauma, and disclosure-related pathways.

Method: At time 1, several measures were used to assess pretrauma variables, trauma variables, and stress reactions upon disclosure for 156 sexually abused children ages 8 to 13 years. At the time 2 follow-up (7 to 36 months following the initial interview), the children were assessed for posttraumatic stress disorder (PTSD) symptoms.

Results: A path analysis involving a series of hierarchically nested ordinary least squares multiple regression analyses indicated three direct paths to PTSD symptoms: avoidant coping, anxiety/arousal, and dissociation, all measured during or immediately after disclosure of sexual abuse. Additionally, age and gender predicted avoidant coping, while life stress and age at abuse onset predicted symptoms of anxiety/arousal. Taken together, these pathways accounted for approximately 57% of the variance in PTSD symptoms.

Conclusions: Symptoms measured at the time of disclosure constitute direct, independent pathways by which sexually abused children are likely to develop later PTSD symptoms. These findings speak to the importance of assessing children during the disclosure of abuse in order to identify those at greatest risk for later PTSD symptoms.

Child sexual abuse is an all too common event in the lives of children and can produce severe psychological damage to victims both at the time of the abuse and years later (1, 2). Many researchers have identified posttraumatic stress disorder (PTSD) as a core manifestation of sexual abuse trauma because of the high frequency with which this disorder and related symptoms appear in sexually abused children (3). A number of retrospective studies of adults have examined the long-term effects of child sexual abuse on later PTSD symptoms (4, 5), but significantly fewer have studied children to examine more immediate PTSD outcomes resulting from child sexual abuse. Even fewer have examined possible mediating mechanisms, such as the children's reactions upon disclosure, in the development of later PTSD symptoms. To our knowledge, the current study is the first to examine pretraumatic vulnerabilities, trauma characteristics, and stress reactions at the time of disclosure as pathways to PTSD symptoms in sexually abused children. Prospective longitudinal designs that assess children shortly after a stressful event, such as this, are ideal for the advancement of the understanding of risk factors for PTSD symptoms (6).

It is important to note that all children in the current study were observed while undergoing a forensic interview, during which the child was asked to recollect and discuss potentially traumatic memories of the sexual abuse. As disclosure of sexual abuse can be an extremely stressful event, responses immediately after such disclosure may offer important information about future risk and resilience. The present study was designed to assess these responses in the service of identifying a model of risk factors for PTSD symptoms in children who disclose sexual abuse.

Research has identified a variety of risk factors for PTSD that are exhibited in the immediate aftermath of a trauma. Dissociation has been found to be particularly important in predicting PTSD (7). Specifically, studies of adults (8–10) and a more recent study of children, reported in this issue of the Journal (11), have indicated that individuals who dissociate either during or soon after a trauma are at greater risk of developing PTSD. Evidence has also suggested that anxiety/arousal responses other than dissociation may predict later PTSD. For example, in a study of adult victims of violent crime, Brewin et al. (12) found that a simple count of reexperiencing symptoms after the trauma independently predicted later PTSD. This finding is consistent with previous research that has linked intrusive symptoms with later problem outcomes, including PTSD (13, 14). Although both dissociative and anxiety/arousal symptoms are included in the diagnosis of acute stress disorder, considerable controversy exists over whether dissociation or anxiety/arousal is more predictive of PTSD (15).

Avoidant coping is a critically important feature of the anxiety/arousal response and has been identified as a risk factor for PTSD (16, 17). Retrospective studies of adult child abuse victims indicate that avoidance strategies such as denial or minimizing are associated with poor psy-
TABLE 1. Pretrauma, Trauma, Disclosure, and Posttrauma Variables for 156 Sexually Abused Children

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretrauma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (male=1, female=0)</td>
<td>0.17</td>
<td>0.38</td>
<td>0.00–1.00</td>
</tr>
<tr>
<td>Age at initial assessment (years)</td>
<td>10.70</td>
<td>1.80</td>
<td>7.92–13.92</td>
</tr>
<tr>
<td>Score for previous life stress, based on parent report of number and severity of stressful events (0=did not happen, 2=major change)</td>
<td>6.38</td>
<td>4.55</td>
<td>0.00–16.00</td>
</tr>
<tr>
<td>Trauma: age at onset of child sexual abuse (years)</td>
<td>8.01</td>
<td>2.67</td>
<td>2.00–13.00</td>
</tr>
<tr>
<td>Disclosure (initial assessment)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoidance (score on avoidance scale from principal-component factor analysis)</td>
<td>0.65</td>
<td>0.63</td>
<td>0.00–3.00</td>
</tr>
<tr>
<td>Dissociation (score on dissociation scale of Trauma Symptom Checklist for Children [22])</td>
<td>8.16</td>
<td>5.04</td>
<td>0.00–20.00</td>
</tr>
<tr>
<td>Anxiety (score on anxiety scale of Trauma Symptom Checklist for Children [22])</td>
<td>7.90</td>
<td>5.26</td>
<td>0.00–24.00</td>
</tr>
<tr>
<td>Posttrauma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interval between initial and follow-up assessments (years)</td>
<td>1.57</td>
<td>0.69</td>
<td>0.58–3.00</td>
</tr>
<tr>
<td>PTSD symptoms at follow-up (score on PTSD scale from Child Behavior Checklist [26])</td>
<td>8.48</td>
<td>5.43</td>
<td>0.00–24.00</td>
</tr>
</tbody>
</table>

Psychological and psychosocial adjustment in adulthood (18, 19). In addition, research on traumatized youth has shown that cognitive suppression and avoidance often lead to problems in multiple domains of functioning (20). Measurement of avoidant coping strategies is complicated by the fact that, by definition, it is difficult to ascertain these symptoms from child self-reports (21). In the current study we surmounted this obstacle by using a behavioral coding system to rate the children’s avoidant behaviors during videotaped forensic interviews.

The current study utilized path analytic techniques to evaluate the ways in which pretrauma variables, trauma characteristics, and stress reactions upon disclosure of sexual abuse are related to later PTSD symptoms in children. As described, anxiety and dissociative reactions in the wake of a trauma have been identified as important components of the acute trauma response, although there is considerable controversy over which is more strongly related to deleterious long-term outcomes. Consequently, a main goal of the current study was to assess the relative importance of dissociation and anxiety in the immediate aftermath of the disclosure of sexual abuse.

Method

Participants and Procedure

The participants in the current study had been referred to a treatment facility that offers services to children considered possible victims of sexual abuse. Each of the participants had been medically examined, interviewed, and videotaped for forensic purposes. The occurrence of sexual abuse was rated by a multidisciplinary evaluation team as confirmed, probable, suspicious, unknown, or “no evidence.” All of the interviews were written up in report form by the interviewers. The inclusion criteria for the current study consisted of the following:

1. The abuse was rated as suspicious, probable, or confirmed.
2. The child was between the ages of 8 and 13 years at the time of the interview.
4. The child was from one of three counties in North Carolina.

Of the possible participants, 156 children met the criteria for the study. The study group comprised 129 girls and 27 boys, with a mean age of 10.7 years (SD=1.8). Of the subjects, 56% were African American, 23% were Caucasian, 12% were Native American, 5% were biracial, and 4% were Hispanic. Abuse was rated as confirmed for 54% of the participants, probable for 18%, and suspicious for 28%.

The information gathered at time 1 came from the clinicians’ written reports, the videotapes, and the Trauma Symptom Checklist for Children (22) collected at the time of the initial forensic interview. Contact with the families for the follow-up at time 2 was initiated through a letter explaining the purpose of the study. After complete description of the study to the subjects, written informed consents were obtained from the parents and informed assents were obtained from the children.

Measures

The following measures, with the exception of the parent’s report of prior life stress and the parent’s report of the child’s PTSD symptoms, were assessed at time 1.

Written reports. In addition to demographic variables (i.e., age, gender), the other variable of interest in the written report following the initial forensic interview was the age at onset of abuse. These items were extracted from the written reports on the children and subsequently coded by the principal investigator (J.B.K.) and a trained research assistant. The reliability estimates were satisfactory, with kappas ranging from 0.59 to 0.90. Discrepant ratings were discussed by the principal investigator and the research assistant until a consensus was reached.

Videotaped interviews. Semistructured investigative interviews were conducted by skilled clinicians trained in the evaluation of sexually abused children, and all interviews were videotaped. A detailed coding system was developed and used to analyze the child’s emotional and behavioral responses throughout the interview, which would be operationalized as child coping variables in the study. The interviews of the participants selected for interrater reliability were randomly sampled intermittently throughout the coding period, and agreement between each of the two coders’ ratings and those of the principal investigator was estimated for 40% of the study participants. The reliability estimates were satisfactory, with kappas ranging from 0.63 to 0.82 and Pearson’s correlation coefficients ranging from 0.41 to 0.98.

Because of the high correlations among many of the items, a principal-component factor analysis was conducted. The most readily interpretable loading pattern was found by using varimax rotation. Each item was assigned to the factor on which it had the highest loading, and no item was allowed to contribute to more than one factor. The scores for the items on each scale were summed and averaged, with higher scores reflecting higher levels of the particular construct. The analysis identified 14 behaviors that factored onto three scales. Of interest to the current study is the avoidance scale, consisting of four items: 1) attempting to distract the interviewer, 2) appearing avoidant (e.g., by not answering questions), 3) appearing distracted, and 4) appearing fidgety (alpha=0.79).

Anxiety and dissociation. The Trauma Symptom Checklist for Children (22) is a 54-item self-report instrument that evaluates
Table 2 shows the correlations among pretrauma, trauma, disclosure, and posttrauma variables for 156 sexually abused children. The table presents Pearson correlation coefficients (r) for various variables, including life stress, age at onset of abuse, avoidance, dissociation, anxiety, interval between assessments, and PTSD symptoms. The table notes that a high percentage of the variance in PTSD symptoms (R^2=0.57). Once these variables were chosen, we began to include antecedent variables that would allow anxiety, avoidance, and dissociation to serve as potential mediators. In order to constrain the number of paths in this model, bivariate relationships between any variable and PTSD that did not reach conventional levels of significance (p<0.05) were deleted. In addition, the model was further constrained by temporal relationships. The final path analytic model provided excellent fit indices (comparative fit index=1.00, Tucker-Lewis index=1.00, root mean square error of approximation=0.00).

**Results**

**Preliminary Analyses**

Table 1 displays descriptive statistics for all variables of interest to the current study. A series of t tests and analyses of variance revealed no significant differences in time 2 PTSD symptoms as a function of gender, race, or county. Table 2 presents Pearson’s correlation coefficients for the associations among pretrauma and trauma variables, disclosure reactions, and later PTSD symptoms.

**Path Analysis**

Figure 1 illustrates the results of the path analysis, with partial correlation coefficients (beta weights) given for each path remaining after nonsignificant paths were removed. The results indicate three direct paths to PTSD symptoms from the disclosure reactions: avoidance, dissociation, and anxiety. They also show four indirect paths to PTSD symptoms from pretrauma and trauma variables: age, gender, life stress, and age at onset of abuse. More specifically, anxiety/arousal served as a mediator between life stress and PTSD and between age at onset and PTSD. In addition, avoidance served as a mediator between age and PTSD and between gender and PTSD. Dissociation was not only a direct predictor of PTSD but also an indirect predictor of PTSD symptoms by way of anxiety symptoms. Taken together, these pathways accounted for 57% of the variance in PTSD. Given the high explanatory value of these pathways, the model helps to shed light on the relatively unknown course of PTSD in sexually abused children.

---

**Table 2. Correlations Among Pretrauma, Trauma, Disclosure, and Posttrauma Variables for 156 Sexually Abused Children**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Correlation (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Life Stress</td>
</tr>
<tr>
<td>Pretrauma</td>
<td></td>
</tr>
<tr>
<td>Age at initial assessment</td>
<td>0.06</td>
</tr>
<tr>
<td>Previous life stress</td>
<td>0.29*</td>
</tr>
<tr>
<td>Trauma: age at onset of child sexual abuse</td>
<td>-0.17</td>
</tr>
<tr>
<td>Disclosure (initial assessment)</td>
<td></td>
</tr>
<tr>
<td>Avoidance</td>
<td>-0.11</td>
</tr>
<tr>
<td>Dissociation</td>
<td>0.68**</td>
</tr>
<tr>
<td>Anxiety</td>
<td>-0.16</td>
</tr>
<tr>
<td>Posttrauma</td>
<td></td>
</tr>
<tr>
<td>Interval between initial and follow-up assessments</td>
<td></td>
</tr>
<tr>
<td>PTSD symptoms at follow-up</td>
<td></td>
</tr>
</tbody>
</table>

*p<0.05.  **p<0.01.
Discussion

To our knowledge, this is the first study to examine the utility of a prospective model incorporating pretrauma, trauma, and disclosure variables in pathways leading to PTSD in sexually abused children. The results of this study demonstrate that sexually abused children who exhibit symptoms of avoidance, anxiety/arousal, or dissociation either during or immediately following disclosure of abuse are at increased risk of developing PTSD symptoms at a later date. These findings speak to the importance of assessing children’s reactions as soon as possible after disclosure.

Given the overwhelming stress experienced by sexual abuse victims during the time of the event as well as during disclosure, it is not surprising that childhood sexual abuse is strongly associated with dissociative responses (7, 32, 33). However, because of previous reliance on retrospective reports of adult sexual abuse victims, the prospective relation between dissociative symptoms and later PTSD symptoms in childhood has remained unclear. This study serves as confirmation that children who report dissociative symptoms immediately after disclosure of abuse are at greater risk for later PTSD symptoms. In fact, dissociation appeared to be the strongest predictor of PTSD symptoms in this group of children. It has been theorized that dissociative responses may prevent the open expression of emotions and cognitions associated with the trauma, which is likely to lead to insufficient processing of the trauma, more reexperiencing symptoms, and consequently, worse PTSD symptoms (34–36).

Dissociation may also represent the phenotype of a biobehavioral vulnerability to traumatic events. A number of researchers have drawn connections between the anxiety/hyperarousal symptoms and the sympathetically mediated fight-or-flight response and between symptoms of dissociation and the parasympathetically mediated “freeze” or “immobilization” response (37–40). Although the current study did not directly measure these biobehavioral systems, i.e., sympathetic and parasympathetic activity, the independent contributions made by the dissociation and anxiety/arousal pathways in predicting PTSD suggest independence of the biobehavioral processes.

Further, as dissociation by this theory is presumed to be a more primitive response that occurs only after the fight-or-flight arousal system has been overwhelmed, it makes sense that there is a unidirectionality of the relationship between dissociation and anxiety/arousal found in this study. Dissociation is significantly associated with anxiety/arousal, whereas anxiety/arousal is not significantly associated with dissociation. In other words, if a child must initially experience arousal/anxiety to reach a dissociated state, then children with dissociation would necessarily experience some symptoms of arousal/anxiety.

Anxiety/arousal symptoms following trauma have been identified in the adult literature as critical in predicting later PTSD (12, 41); however, almost no studies to date have examined the predictive utility of these symptoms in relation to later PTSD in children. The results of the current study demonstrate that anxiety/arousal symptoms immediately after disclosure of sexual abuse make a unique contribution to the development of PTSD symptoms.

It is noteworthy that life stress positively predicted anxiety/arousal symptoms. In addition, age at onset of the abuse negatively predicted anxiety/arousal symptoms, indicating that the earlier the abuse began, the more likely the child was to demonstrate these symptoms at the time of the disclosure. Rutter (42) hypothesized that the accumulation of multiple stressors in children's lives dramatically increases the risk of permanent developmental damage and the manifestation of PTSD symptoms.

Although children’s use of coping strategies has been identified as a particularly promising area of study (43), this field of inquiry has been hampered by the traumatic nature of sexual abuse and its legal and clinical implications. We believe that this study is the first to demonstrate a positive relation between avoidant coping immediately after the discovery of trauma and later PTSD symptoms in sexually abused children. These findings are consistent with studies of adults suggesting that strategies such as denial, minimization, or purposeful forgetting are associated with greater psychological difficulties in the long term (19, 44).

The model also demonstrates that boys are more likely to exhibit avoidant behaviors upon disclosure. This finding is consistent with a number of studies showing higher levels of emotional expression in girls (45, 46) and women (47) than in boys and men, respectively. The current study also indicated that younger children are more likely to be avoidant upon disclosure. Peterson (48) described coping as a process that is significantly influenced by development. Younger children are less likely to have the language

FIGURE 1. Path Analytic Model for the Development of PTSD in 156 Sexually Abused Children*

*The values in the model are partial correlation coefficients (beta weights). R²=0.57.
capacity and emotion identification abilities that older children have, thereby making it more difficult for them to express their thoughts and feelings.

Limitations

Although the current study exhibits important strengths, there are several limitations that should be noted. The findings generalize to the population of children identified by county social service departments as possibly victimized by sexual abuse, but their generalizability to the total population of abused children is unclear. It is also possible that the forensic interview represented the first disclosure of abuse for some children, whereas other children may have disclosed it in great detail to other people prior to the interview. The findings of studies on the relationship between disclosure and psychological functioning have been equivocal (49). Consequently, one must consider the possibility that the child’s reactions observed upon disclosure are mainly a response to the interview itself and are unrelated to the child’s attempt to cope with traumatic reminders.

Clinical Implications

This research has crucial implications for all providers conducting clinical or forensic evaluations and treatments for sexually abused children by helping to identify children who appear to be the most at risk for developing PTSD symptoms. The findings lend support to the notion that children who exhibit avoidance, anxiety/arousal, and especially, dissociation during disclosure are likely to need treatment in order to prevent the development of later PTSD symptoms. Coping is by definition a consciously chosen means of dealing with stress and thereby amenable to modification. Therefore, coping strategies such as avoidance are directly relevant as intervention targets in PTSD treatment (34, 51), the current findings suggest that children who exhibit avoidance, anxiety/arousal, and PTSD symptoms are directly relevant as intervention targets in PTSD treatment in order to prevent the development of later PTSD symptoms. Coping is by definition a consciously chosen means of dealing with stress and thereby amenable to modification. Therefore, coping strategies such as avoidance are directly relevant as intervention targets in therapy (50). Consistent with reports of the outcomes of PTSD treatment (34, 51), the current findings suggest that treatments encouraging open expression of thoughts and feelings surrounding the abuse can help to deter growth in PTSD symptoms over time. Further, the relative independence of dissociative and anxiety/arousal pathways to PTSD symptoms suggests the possibility for targeted treatments aimed at particular biobehavioral systems.

References

PTSD in Sexually Abused Children


47. Sells DJ, Martin RB: Gender and modality differences in experiencing and emotional expression. Canadian J Counseling 2001; 35:176–188


Improving the Rates of Quitting Smoking for Veterans With Posttraumatic Stress Disorder

Miles McFall, Ph.D.
Andrew J. Saxon, M.D.
Charles E. Thompson, M.D.
Dan Yoshimoto, M.A.
Carol Malte, M.S.W.
Kristy Straits-Troster, Ph.D.
Evan Kanter, M.D., Ph.D.
Xiao-Hua Andrew Zhou, Ph.D.
Cynthia M. Dougherty, A.R.N.P., Ph.D.
Bonnie Steele, R.N., Ph.D.

Objective: Smoking is highly prevalent and refractory among people with posttraumatic stress disorder (PTSD). This study aimed to improve the rate of quitting smoking for veterans with PTSD by integrating treatment for nicotine dependence into mental health care.

Method: Smokers undergoing treatment for PTSD (N=66) were randomly assigned to 1) tobacco use treatment delivered by mental health providers and integrated with psychiatric care (integrated care) versus 2) cessation treatment delivered separately from PTSD care by smoking-cessation specialists (usual standard of care). Seven-day point prevalence abstinence was the primary outcome, measured at 2, 4, 6, and 9 months after random assignment. Data were analyzed by using a generalized estimating equations approach following the intent-to-treat principle.

Results: Subjects assigned to integrated care were five times more likely than subjects undergoing the usual standard of care to abstain from smoking across follow-up assessment intervals (odds ratio=5.23). Subjects in the integrated care condition were significantly more likely than subjects in usual standard of care to receive transdermal nicotine and nicotine gum. They also received a greater number of smoking-cessation counseling sessions. Stopping smoking was not associated with worsening symptoms of PTSD or depression.

Conclusions: Smoking-cessation interventions can be safely incorporated into routine mental health care for PTSD and are more effective than treatment delivered separately by a specialized smoking-cessation clinic. Integrating cessation treatment into psychiatric care may have the potential for improving smoking quit rates in other populations of chronically mentally ill smokers.

Posttraumatic stress disorder (PTSD) is one of the most prevalent mental disorders (1, 2), particularly among Veterans Administration (VA) health care enrollees (3, 4). In the general population, PTSD is associated with high rates of smoking (45% [5] versus the national average of 23% [6]) and a fourfold increased risk for nicotine dependence (7). The rates of smoking in veterans with PTSD (53%–66% [8–10]) are approximately double those of VA enrollees in general (30% [11]). Moreover, a greater proportion of smokers with combat-related PTSD smoke heavily (>25 cigarettes per day) compared with veterans who smoke but do not have PTSD (48% versus 28% [8]). PTSD is associated with a smoking quit rate of only 23% (5), about half that of lifetime smokers without a mental disorder (12), and falls third from the bottom in a ranking of quit rates for 13 mental disorders (5). Smokers with PTSD also experience nicotine withdrawal symptoms in response to encounters with trauma-related stimuli (13) and report smoking in order to relieve anxiety and tension (8). Taken together, this research suggests a dynamic relationship between PTSD and tobacco use that argues for a coordinated approach to the treatment of both disorders.

Several pharmacological and behavioral treatments for nicotine dependence have shown efficacy in controlled clinical trials (12, 14). However, these treatments are only as useful as the ability of health care organizations to deliver them effectively to individuals who need them most. Primary care providers only infrequently apply even brief, cost-effective tobacco-cessation interventions to smokers (15–18), despite the fact that 60%–70% of smokers want to quit (6, 11). Nicotine-dependence treatment in patients with mental disorders may be particularly neglected since psychiatric patients receive cessation counseling at only 38% of primary care visits and 12% of visits with a psychiatrist (16). Referral of patients to specialty smoking-cessation clinics is a commonly used alternative to primary care-based delivery of tobacco use treatment. However, the effectiveness of these clinics is compromised by poor patient compliance, with rates of attendance as low as 13%–14% (19, 20) and limited capacity to provide repeated intervention for a chronic, relapsing disorder such as nicotine dependence. These limitations create a situation in which only 17% of smokers in the nation’s largest health care system (the VA) report receiving desired cessation treatment in the previous year (11).

The simultaneous treatment of two or more interwoven disorders by a single provider team has shown promising clinical effectiveness for patients with severe mental ill-
PTSD AND QUITTING SMOKING

TABLE 1. Baseline Characteristics of Subjects Receiving Integrated Care Versus the Usual Standard of Carea

<table>
<thead>
<tr>
<th>Variable</th>
<th>Veterans Receiving Integrated Care</th>
<th>Veterans Receiving Usual Standard of Care</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Age (years)</td>
<td>33</td>
<td>88</td>
</tr>
<tr>
<td>Male sex</td>
<td>29</td>
<td>80</td>
</tr>
<tr>
<td>White race</td>
<td>26</td>
<td>79</td>
</tr>
<tr>
<td>Married</td>
<td>18</td>
<td>55</td>
</tr>
<tr>
<td>Unemployed</td>
<td>27</td>
<td>82</td>
</tr>
<tr>
<td>Receiving Department of Veterans Affairs disability</td>
<td>29</td>
<td>86</td>
</tr>
<tr>
<td>Number of current axis I disorders</td>
<td>33</td>
<td>90</td>
</tr>
<tr>
<td>PTSD Checklist score</td>
<td>33</td>
<td>90</td>
</tr>
<tr>
<td>Beck Depression Inventory score</td>
<td>33</td>
<td>90</td>
</tr>
<tr>
<td>Cigarettes smoked/day</td>
<td>33</td>
<td>90</td>
</tr>
<tr>
<td>Expired carbon monoxide (ppm)</td>
<td>33</td>
<td>90</td>
</tr>
<tr>
<td>Years spent smoking</td>
<td>33</td>
<td>90</td>
</tr>
<tr>
<td>Living with other smokers</td>
<td>33</td>
<td>90</td>
</tr>
<tr>
<td>Confidence in quittingb</td>
<td>33</td>
<td>90</td>
</tr>
</tbody>
</table>

a There were no statistically significant differences between study conditions on baseline characteristics (all p>0.05).
b Ratings were made on a 5-point Likert scale ranging from 0=“not at all” to 4=“extremely.”

ness and comorbid substance use disorders (21, 22). This study accordingly tested an approach for improving the effectiveness of tobacco-cessation service delivery by integrating treatment for smoking into the routine mental health care of patients with PTSD. Specifically, this randomized, controlled clinical trial compared the effectiveness of two different methods for delivering guideline-based smoking-cessation treatment to patients receiving VA mental health care: 1) brief smoking-cessation interventions integrated with ongoing mental health care and delivered by mental health providers (integrated care) versus 2) smoking-cessation interventions delivered separately from mental health care by smoking-cessation specialists (usual standard of care). It was hypothesized that integrated care would result in higher rates of smoking cessation than would usual standard of care. A secondary objective of this study was to determine if smoking cessation was associated with worsening PTSD or depression symptoms. Anticipating this possibility is important, given the high co-occurrence of depression with PTSD and met DSM-IV criteria for PTSD. Exclusion criteria were 1) the presence of unstable axis I disorders, and 3) current substance dependence disorder other than tobacco use. After providing a complete description of the study, written informed consent was obtained. The subjects were then randomly assigned to integrated care (N=33) versus the usual standard of care (N=33) for smoking cessation. The rate of refusal to participate in this study was 3% for help-seeking smokers.

Treatment Conditions

All study subjects received psychotropic medications for PTSD throughout their participation in the study, as prescribed by two PTSD clinic psychiatrists and a nurse practitioner. Each subject also received psychotherapy from an assigned case manager who coordinated their mental health care. Case managers included four psychologists, one social worker, an addictions therapist, and a technician.

Integrated Care: Experimental Condition

The subjects randomly assigned to integrated care received smoking-cessation interventions administered by their assigned PTSD clinic prescriber and case manager. Integrated care was modeled after the brief clinical interventions for primary care practitioners published in the U.S. Public Health Service’s clinical practice guideline titled Treating Tobacco Use and Dependence (14). PTSD clinic staff received approximately 3 hours of training, plus-as-needed consultation in smoking-cessation treatment from the clinic director (M.M.). They then delivered interventions using a treatment manual (available upon request from the first author) that operationalized interventions for each session.

The subjects received smoking-cessation protocol medications (bupropion, transdermal nicotine, nicotine polacrilex gum, and nicotine spray) from the psychiatrist or nurse practitioner managing their pharmacological treatment of PTSD. Prescribers in the integrated care and the usual standard of care conditions agreed to adopt the practice of routinely prescribing bupropion, transdermal nicotine, and nicotine gum or spray in order to standardize the medication protocol across conditions. However, the prescribers were allowed to use their discretion to select only some of these medications, depending on patient preferences and medical contraindications.

The core behavioral counseling components of integrated care, administered by case managers, consisted of 1) the assessment of tobacco use status, abstinence history, and individualized reasons for quitting; 2) education about the health risks of smoking and the benefits of quitting; 3) advice to quit smoking; 4) application of motivational interventions for ambivalent smokers; 5) setting a date to quit; 6) behavioral counseling to help subjects pre-

Method

Subjects

Subjects (N=66) were recruited from the VA Puget Sound Health Care System PTSD clinic, which provides specialized outpatient treatment for chronic PTSD. Table 1 reports demographic and smoking characteristics for the study group as well as psychometric data showing moderate to severe levels of symptoms on the PTSD Checklist (28) and the Beck Depression Inventory (29). Significant differences between study conditions on baseline characteristics (all p>0.05).

Subjects were included if they smoked ≥10 cigarettes per day, expressed a willingness to receive smoking-cessation treatment, and met DSM-IV criteria for PTSD. Exclusion criteria were 1) the use of smokeless tobacco, pipes, or cigars; 2) the presence of unstable axis I disorders, and 3) current substance dependence dis-
pare to quit smoking, coping with smoking urges and barriers to abstinence, and developing problem-solving skills; 7) self-help reading materials; 8) intrasession support and assistance in identifying extrasession social support; and 9) self-directed behavioral methods for reducing anxiety, consisting of a relaxation training tape and written materials on stress management. The protocol required case managers to administer five individual behavioral counseling sessions on a once-weekly basis, plus one follow-up contact (the sessions averaged 20 minutes each). These interventions were rolled into regularly scheduled visits addressing PTSD and comorbid mental disorders for subjects whose treatment plan ordinarily included individual sessions with a case manager. For subjects receiving only group therapy for PTSD, case managers scheduled separate individual smoking-cessation sessions. After delivering the six core behavioral counseling sessions, clinicians used their discretion to periodically assess smoking status and reinstate cessation treatment for subjects who relapsed.

Usual Standard of Care: Comparison Condition

The subjects randomly assigned to the usual standard of care were referred to the VA Puget Sound Health Care System’s Smoking Cessation Clinic. Nursing personnel who staffed this clinic were trained at Mayo Clinic’s Nicotine Dependence Treatment Center and had extensive experience in cessation treatment. This clinic delivered U.S. Public Health Service guideline-adherent cessation treatment (14). Clinic providers used the same algorithm as integrated care providers for prescribing smoking-cessation medications. The subjects attended one group orientation class, followed by individual sessions in which they received medications and behavioral counseling. Unrestricted access to usual standard of care treatment sessions was provided to all subjects. That is, the number of treatment contacts received was determined by clinical recommendations of usual standard of care providers and the preferences and initiative of subjects. The subjects assigned to the usual standard of care condition received absolutely no tobacco-cessation interventions from their PTSD clinic providers.

Measures

Delivery of study treatments. The Consumer Health Informations and Performances Sets, an administrative database accounting for services delivered by the VA, was used to track smoking-cessation medications actually received by the subjects. This database provided information about the proportion of subjects who filled prescriptions for study medications as well as doses for these medications. The number of smoking cessation behavioral counseling sessions delivered to the subjects was extracted from the VA’s computerized patient record system, where clinicians are required to document all patient visits. Adherence of integrated care case managers to the treatment manual was assessed by independent review of the computerized patient record system. These records contained provider documentation of interventions delivered during each visit on specifically prepared charting templates that listed a total of 31 specific protocol interventions prescribed by the manual. These interventions constituted the core behavioral counseling components of integrated care (e.g., “set and recorded a quit date,” “identified strategies for coping with withdrawal symptoms”). After each session, case managers checked off interventions delivered on the charting template and provided a narrative summary of smoking status and treatment progress. The study coordinator (D.Y.) independently reviewed chart notes for all providers and treatment sessions in order to compute clinicians’ self-recorded adherence to protocol interventions as a percentage of the total recommended. This review showed that 80% of behavioral counseling interventions prescribed by the treatment manual were reportedly administered.

Smoking outcomes. Seven-day point prevalence abstinence, verified by expired carbon monoxide (≤10 ppm), was the primary outcome measure used to compare integrated care and the usual standard of care at 2, 4, 6, and 9 months after random assignment. The period after randomization provided a consistent marker across subjects of the initiation of the treatment process. The convention of using an end-of-treatment phase or an initial quit date as a starting point for abstinence measurements was not followed because this study aimed to compare tobacco cessation treatments as they are actually practiced in a clinical setting. Specifically, guideline-adherent treatment for a chronic relapsing condition, such as tobacco use disorder, is a continuous process requiring multiple quit attempts and reaplication of interventions aimed at “recycling” patients who relapse (14, 30, 31). Consistent with expert recommendations (12, 14, 31), study treatments were not defined by a discrete episode of care marked by a fixed endpoint, number of sessions, or single quit date.

Repeated 7-day point prevalence abstinence was used as a proxy for prolonged abstinence, consistent with prior research (32–36). Repeated 7-day point prevalence abstinence was based on three consecutive follow-up assessment intervals (4, 6, and 9 months after random assignment). The 2-month assessment was not included in this computation, allowing subjects a stabilization period to recover from lapses after initial intervention (37). Although repeated 7-day point prevalence abstinence is a less conservative measure than prolonged abstinence, the two measures nevertheless correlate highly (r=0.85–0.94 [38, 39]).

Mental health quality assurance outcomes. The PTSD Checklist (28) and the Beck Depression Inventory (29) were completed by subjects at baseline and 6- and 9-month follow-ups. The subjects rated symptoms that occurred during the prior week. The PTSD Checklist and the Beck Depression Inventory were used to assess changes in mental health symptoms over time in order to determine if smoking cessation was associated with worsening symptoms.

Patient satisfaction outcomes. The subjects in each study condition rated their satisfaction with the amount and quality of smoking-cessation treatment received by using a 4-point Likert scale ([1]=“very dissatisfied” and [4]=“very satisfied”). The ratings were gathered at the terminal (9-month) study assessment.

Data Analysis

Abstinence data were analyzed in an intent-to-treat analysis by using marginal logistic regression with the method of generalized estimating equations. This statistical approach to analyzing smoking-cessation outcome data conforms to recommendations by an expert panel from the Society for Research on Nicotine and Tobacco (39). The generalized estimating equation method has several advantages in that it accounts for the correlation of repeated measurement, uses all available data points, and is sensitive to the pattern of change over time (40). Covariates in the analysis were baseline measures of severity of nicotine dependence (the Fagerstrom Test for Nicotine Dependence [41]) and depression (the Beck Depression Inventory). Depression was included as a covariate because of evidence showing that it may adversely affect smoking-cessation outcomes (42–44).

The effect of smoking cessation on mental health symptoms at 6- and 9-month follow-up assessments was analyzed by computing change from baseline scores for the PTSD Checklist and the Beck Depression Inventory at these assessment intervals. Separate Mann-Whitney U tests were then used to compare these change scores for continuing smokers versus subjects who were abstinent at each assessment interval. All categorical variables (other than abstinence status) were analyzed with chi-square tests. Normally distributed continuous variables were analyzed with Student’s t tests. Nonparametric

TABLE 2. Proportion of Subjects Receiving Smoking-Cessation Medications

<table>
<thead>
<tr>
<th>Medication</th>
<th>Veterans Receiving Smoking-Cessation Meds</th>
<th>Veterans Receiving Smoking-Cessation Meds</th>
<th>χ² (df=1)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Veterans Receiving Smoking-Cessation Meds</td>
<td>Veterans Receiving Smoking-Cessation Meds</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Integrated Care</td>
<td>Usual Standard of Care</td>
<td>Integrated Care</td>
<td>Usual Standard of Care</td>
</tr>
<tr>
<td>Bupropion</td>
<td>20 (60.6%)</td>
<td>16 (48.5%)</td>
<td>0.55</td>
<td>0.46</td>
</tr>
<tr>
<td>Transdermal nicotine</td>
<td>31 (93.9%)</td>
<td>22 (66.7%)</td>
<td>6.13</td>
<td>&lt;0.02</td>
</tr>
<tr>
<td>Nicotine gum</td>
<td>29 (87.9%)</td>
<td>14 (42.4%)</td>
<td>13.08</td>
<td>0.001</td>
</tr>
<tr>
<td>Nicotine spray</td>
<td>1 (3.0%)</td>
<td>3 (9.1%)</td>
<td>0.27</td>
<td>0.61</td>
</tr>
</tbody>
</table>

a Prescriptions written for medications filled by the pharmacy.

Results

Delivery of Study Treatments

Smoking-cessation medications. Table 2 presents the proportion of subjects receiving integrated care and usual standard of care who obtained prescriptions for study medications that were filled by the pharmacy. The subjects randomly assigned to integrated care were significantly more likely than subjects receiving usual standard of care to receive prescriptions for transdermal nicotine and nicotine gum, but the differences between study conditions were not significant for other smoking-cessation medications. Subjects in integrated care also received more types of smoking-cessation medications (i.e., gum, patch, spray, or bupropion) than subjects in usual standard of care (integrated care: mean=2.5, SD=0.7; usual standard of care: mean=1.7, SD=1.0) (z=3.21, p<0.01, two-tailed). Ninety-four percent of the subjects in integrated care filled prescriptions for two or more types of smoking-cessation medications over the course of the study compared to 64% of the subjects in usual standard of care (χ²=7.3, df=1, p<0.01, two-tailed). Table 3 shows that subjects in both conditions who received smoking-cessation medications were given doses and a course of treatment consistent with U.S. Public Health Service practice guidelines (14).

Behavioral counseling. All subjects randomly assigned to integrated care participated in at least one behavioral counseling smoking-cessation session, and 88% of the subjects in usual standard of care had at least one treatment contact. Subjects in integrated care received an average of 5.15 (SD=1.2) manual-driven protocol smoking-cessation sessions, compared to an average of 2.6 (SD=2.1) sessions for usual standard of care (z=5.35, p<0.0001, two-tailed). Additional provider-initiated follow-up contacts (mean=3.9, SD=3.9) were delivered to integrated care subjects, involving assessment of smoking status and reapplication of interventions deemed appropriate by the case manager. The number of smoking-cessation sessions received by all subjects was significantly correlated with the number of assessment occasions (four possible) when they were abstinent from smoking (rₛ=0.35, p<0.004, two-tailed).

Smoking-Cessation Outcomes

Subject compliance with follow-up assessments was 83% across all four assessment intervals (83% at month 2, 81% at month 4, 81% at month 6, and 84% at month 9). Medical records and/or retrospective subject interview data were available to determine smoking status for 96% of the remaining observations. Four percent of the observations were missing and were handled by recording subjects as presumptive smokers. The two study conditions did not differ significantly in the number of missing observations (p>0.05).

Figure 1 presents smoking abstinence data for each assessment interval and the results of the generalized estimating equation analysis. At each assessment interval, the odds of not smoking at that interval were over five times greater for the subjects in integrated care than the subjects in usual standard of care (odds ratio=5.23, 95% confidence interval=1.76–15.54, p<0.002, two-tailed). The proportion of subjects who achieved abstinence at one or more assessment intervals was greater for the subjects in integrated care (integrated care=52% versus usual standard of care=25%) (χ²=4.82, df=1, p<0.02, two-tailed). The subjects in integrated care were also abstinent at more follow-up assessment periods (four possible) compared to the subjects in usual standard of care (mean=1.1, SD=1.4, for integrated care versus mean=0.4, SD=0.8, for usual standard of care) (z=2.38, p<0.02, two-tailed). The repeated 7-day point prevalence abstinence rate was 12% for integrated care and 3% for the usual standard of care, a statistically nonsignificant difference (χ²=1.66, df=1, p=0.20, two-tailed).

Quality Assurance Outcomes

For the group as a whole, scores on the PTSD Checklist and the Beck Depression Inventory gathered at 6 and 9 months did not change significantly from baseline (all p>0.05). Change scores were not significantly different for abstainers versus continued smokers at either assessment interval (all p>0.05).

Satisfaction With Treatment

The subjects in integrated care were significantly more satisfied with the amount of smoking-cessation treatment they received compared to the subjects in usual standard of care (mean=3.9, SD=0.3, for integrated care versus mean=3.5, SD=0.7, for the usual standard of care) (z=3.21, p<0.001, two-tailed). Ratings for the quality of treatment were also significantly higher for the integrated care condition (mean=3.7, SD=0.5, for integrated care versus mean=3.4, SD=0.6, for the usual standard of care) (z=2.06, p<0.04, two-tailed).
Transdermal nicotine was a more effective vehicle than the usual standard of care and were significantly more likely to receive transdermal nicotine and nicotine gum than specific smoking-cessation counseling sessions received than usual standard of care, may have improved access to ongoing monitoring and provider-initiated relapse management. Other research has shown a “dose-response” relationship between the number of smoking-cessation contacts and smoking outcomes (14, 33, 35), as well as higher quit rates for extended versus brief tobacco use treatments (45). In fact, some treatment trials suggest that cessation may be related more to the number of counseling sessions received than specific therapeutic methods used for smokers with a positive history of depression (32–35). Successful smoking-cessation treatment for veterans with PTSD may similarly require a greater number of contacts over an extended time.

Guideline-recommended smoking-cessation treatments yielded point prevalence quit rates of 15% to 25% in U.S. Public Health Service meta-analyses of 6,000 studies involving follow-up assessments of at least 5 months (12). Individuals with mental disorders were typically excluded from smoking-cessation clinical trials reviewed by the U.S. Public Health Service, as active psychiatric comorbidities complicate cessation efforts and are associated with significantly reduced quit rates (5, 46–49). The fact that point prevalence quit rates observed in this study fell within the range of those reported in meta-analyses of treatment trials is encouraging, given that our subjects were chronically mentally ill and experienced moderate to severe symptoms of PTSD and depression. Research is clearly needed that improves upon the repeated 7-day point prevalence abstinence quit rates measured at 4, 6, and 9 months after treatment onset were 15%.

Subjects in integrated care participated in a greater number of smoking-cessation counseling sessions than subjects in usual standard of care and were significantly more likely to receive transdermal nicotine and nicotine gum. They were also more likely to receive combination tobacco-cessation pharmacotherapy, which has been shown to improve smoking quit rates over use of monotherapy (14). These findings suggest that integrated care was a more effective vehicle than the usual standard of care for delivering cessation treatments of sufficient intensity. Integrated care may have been more successful at engaging subjects in cessation treatment, as indicated by higher satisfaction ratings for integrated care than usual standard of care and encouraging medication compliance. Additionally, the continuous therapeutic relationship afforded by integrated care, as opposed to the episodic care provided by the usual standard of care, may have improved access to ongoing monitoring and provider-initiated relapse management. Other research has shown a “dose-response” relationship between the number of smoking-cessation contacts and smoking outcomes (14, 33, 35), as well as higher quit rates for extended versus brief tobacco use treatments (45). In fact, some treatment trials suggest that cessation may be related more to the number of counseling sessions received than specific therapeutic methods used for smokers with a positive history of depression (32–35). Successful smoking-cessation treatment for veterans with PTSD may similarly require a greater number of contacts over an extended time.

Discussion

This study demonstrated the feasibility of training mental health providers to integrate guideline-based smoking-cessation treatment into mental health care for veterans with PTSD. PTSD clinic prescribers readily incorporated the delivery of tobacco-cessation medications into their clinical practice. Case managers also documented a high rate of adherence to prescribed behavioral counseling interventions, although the absence of verification through independent ratings of session content limits the validity of this finding. The integrated model of smoking-cessation treatment tested here was more effective for PTSD patients than for care provided by VA smoking-cessation specialists, as measured by point-prevalence abstinence. The difference between study conditions on the repeated 7-day point prevalence abstinence measure (12% for integrated care, 3% for the usual standard of care) was not significant, owing to the limited statistical power of the study. Since completing this randomized trial, PTSD clinic staff have delivered integrated care to 107 additional patients.

Bupropion

Table 3: Quantity of Smoking-Cessation Medications Received by Subjects

<table>
<thead>
<tr>
<th>Medication and Measure</th>
<th>Veterans Receiving Integrated Care</th>
<th>Veterans Receiving Usual Standard of Care</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Bupropion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days’ supply</td>
<td>20</td>
<td>141.8</td>
<td>114.5</td>
</tr>
<tr>
<td>Number of prescriptions filled</td>
<td>20</td>
<td>3.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Nicotine gum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dose (mg/day)</td>
<td>31</td>
<td>23.7</td>
<td>5.4</td>
</tr>
<tr>
<td>Days’ supply</td>
<td>31</td>
<td>103.0</td>
<td>119.1</td>
</tr>
<tr>
<td>Number of prescriptions filled</td>
<td>31</td>
<td>4.6</td>
<td>4.8</td>
</tr>
<tr>
<td>Nicotine spray</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dose (mg/day)</td>
<td>29</td>
<td>7.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Days’ supply</td>
<td>29</td>
<td>79.9</td>
<td>69.5</td>
</tr>
<tr>
<td>Nicotine spray</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of bottles received</td>
<td>1</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Days’ supply</td>
<td>1</td>
<td>14.0</td>
<td></td>
</tr>
<tr>
<td>Number of prescriptions filled</td>
<td>1</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

a Prescriptions written for medications that were filled by the pharmacy.

b Significance testing was not conducted because of the small number of observations in the two study conditions.

This study demonstrated the feasibility of training mental health providers to integrate guideline-based smoking-cessation treatment into mental health care for veterans with PTSD. PTSD clinic prescribers readily incorporated the delivery of tobacco-cessation medications into their clinical practice. Case managers also documented a high rate of adherence to prescribed behavioral counseling interventions, although the absence of verification through independent ratings of session content limits the validity of this finding. The integrated model of smoking-cessation treatment tested here was more effective for PTSD patients than for care provided by VA smoking-cessation specialists, as measured by point-prevalence abstinence. The difference between study conditions on the repeated 7-day point prevalence abstinence measure (12% for integrated care, 3% for the usual standard of care) was not significant, owing to the limited statistical power of the study. Since completing this randomized trial, PTSD clinic staff have delivered integrated care to 107 additional patients.

Bupropion

Table 3: Quantity of Smoking-Cessation Medications Received by Subjects

<table>
<thead>
<tr>
<th>Medication and Measure</th>
<th>Veterans Receiving Integrated Care</th>
<th>Veterans Receiving Usual Standard of Care</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Bupropion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days’ supply</td>
<td>20</td>
<td>141.8</td>
<td>114.5</td>
</tr>
<tr>
<td>Number of prescriptions filled</td>
<td>20</td>
<td>3.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Nicotine gum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dose (mg/day)</td>
<td>31</td>
<td>23.7</td>
<td>5.4</td>
</tr>
<tr>
<td>Days’ supply</td>
<td>31</td>
<td>103.0</td>
<td>119.1</td>
</tr>
<tr>
<td>Number of prescriptions filled</td>
<td>31</td>
<td>4.6</td>
<td>4.8</td>
</tr>
<tr>
<td>Nicotine spray</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dose (mg/day)</td>
<td>29</td>
<td>7.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Days’ supply</td>
<td>29</td>
<td>79.9</td>
<td>69.5</td>
</tr>
<tr>
<td>Nicotine spray</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of bottles received</td>
<td>1</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Days’ supply</td>
<td>1</td>
<td>14.0</td>
<td></td>
</tr>
<tr>
<td>Number of prescriptions filled</td>
<td>1</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

a Prescriptions written for medications that were filled by the pharmacy.

b Significance testing was not conducted because of the small number of observations in the two study conditions.
interventions in an effort to prevent early relapse to smoking. Application of integrated care on a systemwide basis, even with a prolonged smoking quit rate of 12%, still has a potential for producing 10,000 quitters from within the population of PTSD patients seen at the VA alone (estimates derived from VA databases [11, 50, 51]).

There exist remarkably few smoking-cessation clinical trials involving individuals with current mental disorders that provide a standard of comparison for the present study. A number of studies have been conducted with smokers having a history of past major depression (32–36, 52, 53). However, these studies do not provide the appropriate standard of comparison for the present study, and the preponderance of evidence from these trials does not substantiate the hypothesis that history-positive smokers are less likely to quit than history-negative smokers (54). Clinical trials involving smokers with current mental disorders who meet the minimum recommended standard for abstinence at follow-up assessments (≥6 months [37, 55]) have focused on schizophrenia and alcohol dependence. (Single point prevalence abstinence rates at the terminal 6-month follow-up assessment for studies of smokers with schizophrenia: 11% [56], 12% [57], 7%–17% [58], and 19% [59]; single point prevalence abstinence rates at the terminal 12-month assessment for studies of alcohol-dependent patients: 12% [60] and 13%–19% [61].) Quit rates from these studies of clinical samples are roughly equivalent to, if not slightly lower than, the single-point prevalence quit rates for integrated care reported here. The only study (61) reporting prolonged abstinence (repeated 7-day point prevalence abstinence at follow-up months 1, 3, 6, and 12) found quit rates of 10% and 12% for two active tobacco-cessation treatments involving recovering alcoholics. These results are comparable to the repeated 7-day point prevalence abstinence quit rate of 12% in the present study.

The rate of smoking abstinence for integrated care subjects at study termination was less than half that observed at the 2-month assessment. Several possible mechanisms that may contribute to smoking relapse in PTSD patients deserve further exploration. Specifically, pre- and post-cessation negative affective states (24, 32, 33, 44, 46, 62) and increased symptoms of depression immediately following a quit attempt (63) increase relapse susceptibility and may have suppressed sustained quit rates in this study. Heightened withdrawal sensitivity has predicted relapse in smokers with a history of depression in some studies (64, 65) and may possibly explain continued smoking in PTSD patients as well.

Stopping smoking in this treatment trial was not associated with worsening symptoms of PTSD or depression measured at 6 and 9 months after treatment onset. Our results are more consistent with studies showing that mood disturbances do not result from stopping smoking (32, 46, 66) than with studies showing that they do (25–27, 67) among individuals with a history of depression. However, comparing our results with those of others is attenuated because of the marked differences in study groups (i.e., current PTSD versus a history of major depression). Three other possible explanations for this finding also deserve consideration. First, transient symptom exacerbations related to postcessation nicotine withdrawal may have occurred shortly after subjects’ cessation attempts but were undetected because outcomes were only measured at four discrete time points in our study. Second, the fact that 86% of the subjects in the present research were receiving antidepressant medications (other than bupropion for smoking cessation) for PTSD may have prevented potential deterioration effects. This possibility is supported by evidence showing that sertraline reduces nicotine withdrawal symptoms (53) and that nortriptyline alleviates postcessation negative affect (35) in smokers with a history of major depression. Third, subjects who experienced worsening psychiatric symptoms upon trying to quit smoking may have simply resumed smoking in order to manage these symptoms. Our data do not support this latter explanation since symptoms of PTSD and depression did not change from baseline to the 6- and 9-month assessment for subjects who continued to smoke compared to those who quit. However, the nonsignificant difference between subjects who quit and those who did not on symptom measures may reflect low statistical power owing to the small number of quitters involved in comparisons at months 6 and 9.

Interpretation of this study’s findings may be limited by several methodological shortcomings. This investigation did not follow the convention of measuring outcomes from a clearly demarcated quit date or the end of the intervention period because smoking-cessation treatment was conceptualized as a continuous process involving reappli-
cation of treatment and multiple quit attempts (31). The comparability of findings with other research that measures outcomes from an initial quit date may therefore be limited. Biomarkers of longer-term nonsmoking status were not used but should ideally have supplemented expired carbon monoxide in order to verify subject self-reports. Continuous abstinence was also not assessed because it is not verifiable with standard biological assays. However, continuous abstinence data based on subject self-report alone would have been informative because subject reports of smoking status are generally valid (68).

Proportions of subjects who received prescriptions filled by the pharmacy were reported, but a specific measure of medication adherence was not included. Thus, it is impossible to determine to what degree, if any, medication adherence may have affected outcomes. PTSD patients with a current substance abuse disorder were included in the study, but those with current substance dependence were not. Thus, our findings may not generalize to smokers with PTSD and comorbid substance dependence disorders. Finally, a much larger study group is clearly required in order to conclude definitively that integrated care is more effective than the usual standard of care for smoking cessation in other clinical settings.

This service delivery study was conducted in the spirit of a clinical effectiveness trial in order to promote generalization to “real world” health care settings (69). To this end, treatments were represented as they are likely to be actually practiced, and broad subject inclusion criteria were used to ensure that the study group reflected the population of VA PTSD clinic patients. Despite threats to internal validity inherent in such trials (69), this study provided a favorable test of the feasibility and outcome of APA recommendations (49) to implement smoking-cessation treatment in practice-based settings for patients with current mental disorders. Additional studies of integrated models of smoking-cessation treatment for mentally ill patients are warranted, given the extraordinarily high prevalence of heavy smoking in these individuals (41% [5] to 50% [70]).

References

1. Narrow WE, Rae DS, Robins LN, Regier DA: Revised prevalence estimates of mental disorders in the United States. Arch Gen Psychiatry 2002; 59:115–123

Presented in part at the International Society for Traumatic Stress Studies, New Orleans, Dec. 6–9, 2001. From the Northwest Network Mental Illness Research, Education, and Clinical Center, the Center for Excellence in Substance Abuse Treatment and Education, and the Biological Statistics Unit of the Health Services Research and Development Center of Excellence, Department of Veterans Affairs, Seattle, Wash.; and the University of Washington School of Medicine, the Department of Psychiatry and Behavioral Sciences, and the School of Nursing, Seattle. Address correspondence and reprint requests to Dr. McFall, VA Puget Sound Health Care System, PTSD Programs (S-116 MHC), 1660 S. Columbian Way, Seattle, WA 98108; miles.mcfall@med.va.gov (e-mail).

Supported by the Department of Veterans Affairs Northwest Network Mental Illness Research, Education, and Clinical Center, the Center for Excellence in Substance Abuse Treatment and Education, and a grant from the University of Washington Alcohol and Drug Abuse Institute.
PTSD AND QUITTING SMOKING


36. Hall SM, Humfleet GL, Reus VI, Munoz RF, Hartz DT, Maude-Griffin R: Psychological intervention and antidepressant treatment in smoking cessation. Arch Gen Psychiatry 2002; 59:930–936


42. Anda RF, Williamson DF, Escobedo LG, Mast EE, Giovino GA, Remington PL: Depression and the dynamics of smoking. JAMA 1990; 264:1541–1545


http://ajp.psychiatryonline.org

Am J Psychiatry 162:7, July 2005


64. Glassman AH, Stetner F, Walsh BT, Raizman PS, Fleiss JL, Cooper TB, Covey LS: Heavy smokers, smoking cessation, and clonidine. JAMA 1988; 259:2863–2866


Longitudinal Course of Posttraumatic Stress Disorder and Posttraumatic Stress Disorder Symptoms in a Community Sample of Adolescents and Young Adults

Axel Perkonigg, Ph.D.
Hildegard Pfister, Dipl.-Inf.
Murray B. Stein, M.D., M.P.H.
Michael Höfler, Dipl.-Stat.
Roselind Lieb, Ph.D.
Andreas Maercker, M.D., Ph.D.
Hans-Ulrich Wittchen, Ph.D.

Objective: Few studies have focused on the natural course of posttraumatic stress disorder (PTSD) and its determinants in samples of the general population. The authors examined determinants of remission and chronicity of PTSD and associations with other disorders in a prospective community sample.

Method: The data were drawn from a prospective, longitudinal epidemiological study of adolescents and young adults (age 14–24 years) in Munich, Germany (N=2,548). The course of PTSD from baseline to follow-up 34–50 months later was studied in 125 respondents with DSM-IV PTSD or subthreshold PTSD at baseline.

Results: Although 52% of the PTSD cases remitted during the follow-up period, 48% showed no significant remission of PTSD symptoms. Respondents with a chronic course were more likely to experience new traumatic event(s) during follow-up (odds ratio=5.21, 95% confidence interval [CI]=1.95–13.92), to have higher rates of avoidant symptoms at baseline (odds ratio=10.16, 95% CI=1.73–59.51), and to report more help seeking (odds ratio=5.50, 95% CI=1.04–29.05), compared to respondents with remission. Rates of incident somatoform disorder (odds ratio=4.24, 95% CI=1.60–11.19) and other anxiety disorders (odds ratio=4.07, 95% CI=1.15–14.37) were also significantly associated with a chronic course.

Conclusions: PTSD is often a persistent and chronic disorder. Specific symptom clusters—especially avoidant symptoms—might be associated with the course of PTSD. In addition, the occurrence of new traumatic events differentiates PTSD cases with a chronic course from those with remission.

Posttraumatic stress disorder (PTSD) is a prevalent disorder for which overall community lifetime prevalence estimates range from a minimum of 1% in earlier DSM-III studies to a maximum of 12.3% in more recent surveys (1–8).

A few epidemiological studies, most of them retrospective, have focused on the natural course of PTSD and its determinants. Chronic DSM-III or DSM-III-R PTSD was frequently reported in rape victims, victims of torture and political violence, refugees, and combat veterans (9–11). Kilpatrick et al. (9) reported that 16.5% of women who had been raped continued to meet PTSD diagnostic criteria an average of 17 years afterward. Risk factors for lifetime chronic (duration at least 1 year) PTSD have been reported for a random sample of young adults in the Detroit metropolitan area by Breslau and Davis (10). Compared to young adults with nonchronic PTSD, those with chronic PTSD had a higher number of DSM-III-R PTSD symptoms and higher rates of interpersonal numbing and overreactivity to stimuli that symbolized the stressor, as well as higher rates of psychiatric comorbidity and other medical conditions. Davidson et al. (11) noted significant differences between chronic and acute DSM-III PTSD in a general population sample; chronic PTSD was associated with reduced social support, a greater frequency of social phobia, and greater avoidant symptoms.

Most of our knowledge of the course of PTSD and its determinants is based on prospective (12–15) or retrospective (16) cohort studies involving people who have experienced specific types of traumatic events (e.g., natural disasters). For example, a prospective study among Australian firefighters reported a 30% rate of PTSD identified by assessment with the General Health Questionnaire 29 months after a bush fire (14). Retrospective data from adult survivors of the Buffalo Creek flood showed that after 14 years, 28% of the survivors with PTSD still had not had remission (15). Results from the National Vietnam Veterans Readjustment Study showed that 15% of all male Vietnam veterans still had PTSD 19 years after combat exposure (16). Some studies found that specific symptom clusters and clinical features fluctuated over time and sometimes increased in intensity after several decades of decline (14). Furthermore, numerous pre-, peri-, and post-exposure risk factors have been reported, suggesting a multifactorial model of longitudinal course with various pathways (14). For example, studies of acute stress reactions or acute stress disorder provided support for associations between the immediate response after a traumatic event and PTSD, but there is still much to learn about the course of PTSD in the general population.
event (i.e., acute motor restlessness) and the longitudinal course of PTSD (17). Some studies suggested a significant role of cognitive processing style (18), coping behavior (13), and social support (19). However, generalizability of these findings is an unresolved issue.

Several epidemiological studies have shown that PTSD is strongly associated with comorbid disorders (1, 5, 20). Most of these findings are based on retrospective or cross-sectional data, and the causal explanations of psychiatric sequelae or temporal associations between PTSD and other disorders are unclear (21, 22). The few extant studies on the course of PTSD and psychiatric risk factors showed that other anxiety disorders and depressive symptoms after the onset of PTSD are related to a more chronic course (12, 23). However, no studies with an epidemiological and longitudinal approach have assessed a broader range of psychiatric risk factors after various types of traumatic events in the general population.

The high degree of comorbidity was also confirmed in a community sample of adolescents and young adults ages 14–24 years in the Early Developmental Stages of Psychopathology study (5). We previously reported that 1.3% of the baseline sample (0.4% of male respondents and 2.2% of female respondents) and approximately 8% of those who reported traumatic events fulfilled the DSM-IV criteria for PTSD. PTSD and traumatic events were strongly associated with other DSM-IV disorders that occurred as either a primary or secondary disorder before or after the onset of PTSD. We also previously reported the patterns of incidence of PTSD in this community sample (24).

In this article, we report results from the follow-up component of the Early Developmental Stages of Psychopathology study, which focused on the 42-month longitudinal course of PTSD and PTSD symptoms. We explored the following questions:

1. How many persons with subthreshold PTSD or full PTSD at baseline had entire or partial remission of PTSD during the follow-up period?
2. How do persons with a chronic course differ from persons who experience remission(s) during follow-up with respect to specific PTSD characteristics and risk factors?
3. Is a chronic course of PTSD associated with the onset of other disorders?

### Method

#### Sample and Overall Design

The data presented here were collected as part of the Early Developmental Stages of Psychopathology study (25, 26). The Early Developmental Stages of Psychopathology study was designed as a prospective-longitudinal survey to explore prevalence and incidence, familial and other risk factors, and comorbidity and course of mental disorders in a population sample of adolescents and young adults randomly and proportionally drawn from regional registries to represent the distribution of persons age 14–24 years.

#### TABLE 1. Status at Baseline and 34–50-Month Follow-Up and Course From Baseline to Follow-Up in Respondents With Subthreshold Posttraumatic Stress Disorder (PTSD) and Full DSM-IV PTSD at Baseline in a Community Sample of Adolescents and Young Adults Age 14–24 Years

<table>
<thead>
<tr>
<th>Status and Course</th>
<th>Respondents With Subthreshold PTSD at Baseline (N=101)</th>
<th>Respondents With Full DSM-IV PTSD at Baseline (N=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline status</td>
<td>N</td>
<td>Conditional Probability (Weighted %)</td>
</tr>
<tr>
<td>All respondents (N=2,548)</td>
<td>101</td>
<td>4.4c</td>
</tr>
<tr>
<td>Trauma victims (N=393)</td>
<td>101</td>
<td>26.1d</td>
</tr>
<tr>
<td>Male respondents</td>
<td>43</td>
<td>44.8</td>
</tr>
<tr>
<td>Female respondents</td>
<td>58</td>
<td>55.2</td>
</tr>
<tr>
<td>Follow-up status (34–50 months)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTSD symptoms</td>
<td>20</td>
<td>21.1</td>
</tr>
<tr>
<td>Subthreshold PTSD</td>
<td>20</td>
<td>18.3</td>
</tr>
<tr>
<td>Full DSM-IV PTSD</td>
<td>4</td>
<td>4.2</td>
</tr>
<tr>
<td>Other</td>
<td>57</td>
<td>56.7</td>
</tr>
<tr>
<td>Course from baseline to follow-up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remission</td>
<td>57</td>
<td>56.7</td>
</tr>
<tr>
<td>Chronic course</td>
<td>44</td>
<td>43.3</td>
</tr>
<tr>
<td>Symptomatic (partial remission)</td>
<td>20</td>
<td>21.1</td>
</tr>
<tr>
<td>Subthreshold or full DSM-IV PTSD</td>
<td>24</td>
<td>22.2</td>
</tr>
<tr>
<td>New traumatic events after baseline</td>
<td>44</td>
<td>44.5</td>
</tr>
</tbody>
</table>

a Respondents with subthreshold PTSD fulfilled the A (traumatic event, fear), B (persistent reexperiences), and E (duration) criteria for DSM-IV PTSD but did not fulfill completely the C (avoidance or numbing of general responsiveness) and/or D (increased arousal) criteria, although they reported at least one symptom in each of the C and D criteria categories with a duration of more than 1 month.

b Probability among respondents with subthreshold PTSD or DSM-IV PTSD.

c Prevalence in the entire study sample.

d Rate among trauma victims.

e Respondents with PTSD symptoms fulfilled the A and the B criteria for DSM-IV PTSD but did not meet the criteria for subthreshold PTSD or the full criteria for DSM-IV PTSD.

f Symptoms may be present, but they do not fulfill the criteria for PTSD symptoms, subthreshold PTSD, or DSM-IV PTSD (i.e., duration and persistence criteria).
TABLE 2. Characteristics of Respondents With Posttraumatic Stress Disorder (PTSD) at Baseline and With Complete Remission or a Chronic Course at 34–50-Month Follow-Up in a Community Sample of Adolescents and Young Adults Age 14–24 Years

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Respondents With PTSD Remission at Follow-Up (N=65)</th>
<th>Respondents With a Chronic Course of PTSD at Follow-Up (N=60)</th>
<th>Analysisb</th>
<th>Respondents Without New Traumas (N=74)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weighted %</td>
<td>Weighted %</td>
<td>Odds Ratio</td>
<td>95% CI</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>31</td>
<td>65.1</td>
<td>16</td>
<td>34.9</td>
</tr>
<tr>
<td>Female</td>
<td>34</td>
<td>46.5</td>
<td>44</td>
<td>53.5</td>
</tr>
<tr>
<td>Age group at baseline (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14–17</td>
<td>18</td>
<td>57.6</td>
<td>12</td>
<td>42.4</td>
</tr>
<tr>
<td>18–24</td>
<td>47</td>
<td>53.0</td>
<td>48</td>
<td>47.1</td>
</tr>
<tr>
<td>Trauma type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonassaultive</td>
<td>25</td>
<td>55.4</td>
<td>21</td>
<td>4v4.6</td>
</tr>
<tr>
<td>Assaultive</td>
<td>40</td>
<td>52.6</td>
<td>39</td>
<td>47.4</td>
</tr>
<tr>
<td>Age at onset of baseline trauma (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥12</td>
<td>43</td>
<td>56.1</td>
<td>37</td>
<td>43.9</td>
</tr>
<tr>
<td>&lt;12</td>
<td>22</td>
<td>48.6</td>
<td>23</td>
<td>51.4</td>
</tr>
<tr>
<td>Number of traumas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>52</td>
<td>57.4</td>
<td>42</td>
<td>42.6</td>
</tr>
<tr>
<td>&gt;1</td>
<td>13</td>
<td>42.1</td>
<td>18</td>
<td>58.0</td>
</tr>
<tr>
<td>Number of symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥5</td>
<td>26</td>
<td>56.3</td>
<td>17</td>
<td>43.7</td>
</tr>
<tr>
<td>&gt;5</td>
<td>39</td>
<td>52.4</td>
<td>43</td>
<td>47.6</td>
</tr>
<tr>
<td>Number of avoidant symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2</td>
<td>42</td>
<td>62.2</td>
<td>26</td>
<td>37.8</td>
</tr>
<tr>
<td>≥2</td>
<td>23</td>
<td>44.2</td>
<td>34</td>
<td>55.8</td>
</tr>
<tr>
<td>Number of depressive symptoms and interpersonal numbing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2</td>
<td>54</td>
<td>57.5</td>
<td>38</td>
<td>42.5</td>
</tr>
<tr>
<td>≥2</td>
<td>11</td>
<td>44.0</td>
<td>22</td>
<td>56.0</td>
</tr>
<tr>
<td>Number of arousal symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2</td>
<td>19</td>
<td>59.9</td>
<td>12</td>
<td>40.2</td>
</tr>
<tr>
<td>≥2</td>
<td>46</td>
<td>51.8</td>
<td>48</td>
<td>48.2</td>
</tr>
<tr>
<td>Impairment due to symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes, moderate</td>
<td>53</td>
<td>58.7</td>
<td>40</td>
<td>41.3</td>
</tr>
<tr>
<td>Yes, severe/very severe</td>
<td>12</td>
<td>41.1</td>
<td>20</td>
<td>58.9</td>
</tr>
<tr>
<td>Help seeking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>54</td>
<td>59.7</td>
<td>37</td>
<td>40.3</td>
</tr>
<tr>
<td>Yes</td>
<td>11</td>
<td>38.1</td>
<td>23</td>
<td>61.9</td>
</tr>
<tr>
<td>New traumas between baseline and follow-up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>50</td>
<td>66.9</td>
<td>24</td>
<td>33.1</td>
</tr>
<tr>
<td>Yes</td>
<td>15</td>
<td>33.3</td>
<td>36</td>
<td>66.7</td>
</tr>
<tr>
<td>Presence of any other baseline mental disorderd</td>
<td>37</td>
<td>55.8</td>
<td>34</td>
<td>44.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.8</td>
<td>3.0</td>
<td>1.57</td>
<td>0.97–2.54</td>
</tr>
</tbody>
</table>

a Includes respondents with both subthreshold PTSD and full DSM-IV PTSD. Respondents with subthreshold PTSD fulfilled the A, B (traumatic event, fear, and persistent reexperiences), and E criteria (duration) for DSM-IV PTSD but did not fulfill completely the C (avoidance or numbing of general responsiveness) and/or D criteria (increased arousal), although they reported at least one symptom in each of the C and D criteria categories with a duration of more than 1 month.

b Multiple logistic regression analyses of variables related to chronic course.

c Significant difference between respondents with remission and respondents with a chronic course.

d Includes substance use disorders, mood disorders, other anxiety disorders, somatoform disorders, and eating disorders.

e Standardized scores are reported; higher scores are associated with lower self-competence.

years in the area. After complete description of the study to the subjects, written informed consent was obtained.

The study was divided into three waves. The first wave was conducted in 1995 (baseline) and included the representative sample of persons age 14–24 years (N=3,021). The second wave was conducted in 1996–1997 (time 1) and included only respondents who were age 14–17 years at baseline (N=1,228). The third wave was conducted in 1998–1999 (time 2) and again included all respondents who were age 14–24 years at baseline (N=2,548). Detailed information on the sampling procedure has been reported elsewhere (25). Briefly, at baseline, a total of 3,021 interviews were completed (response rate=71%). The first follow-up (time 1), among the respondents who were age 14–17 years at baseline, was conducted an average of 19.7 months (range=13–25.6) after the baseline interviews (response rate=88%). The second follow-up (time 2), which was intended to include all baseline participants, was conducted in 1998–1999, an average of 42 months (range=34–50) after the baseline interviews (response rate=84%). The findings reported here are based on baseline and second follow-up data (N=2,548). For the younger cohort (respondents age 14–17 years at baseline), information from the first follow-up (i.e., on new traumatic events) was added.
The sociodemographic characteristics of the baseline and entire follow-up (baseline to time 2) sample have been published (5, 26). Briefly, at baseline, most of the respondents were attending school (89%) and living with their parents (97.8%). About 10% were in job training. The majority was classified as belonging to the middle class (61.4%). Noteworthy changes in sociodemographic characteristics from baseline to the second follow-up were found for school status (42% were attending school at follow-up) and employment status (24% were in a job training program and 12% were employed at follow-up).

Diagnostic Assessment

Diagnostic assessments at the baseline and the two follow-up investigations were based on the computer-assisted personal version of the Munich-Composite International Diagnostic Interview (27), which allows for the assessment of symptoms, syndromes, and diagnoses of 48 mental disorders according to the DSM-IV criteria and for collection of data on onset, duration, severity, and psychosocial impairment. Diagnostic findings were obtained by using the Munich-Composite International Diagnostic Interview/DSM-IV algorithms. At baseline, the lifetime version of the Munich-Composite International Diagnostic Interview was used. At each follow-up, the interval version was applied. In all assessments, the Munich-Composite International Diagnostic Interview was supplemented by a separate respondents’ booklet that included several scales and questionnaires for assessing psychological constructs relevant to the study (26). For the purpose of this examination, we additionally used a self-competence scale (28). Test-retest reliability and validity of the full Munich-Composite International Diagnostic Interview, along with descriptions of the Munich-Composite International Diagnostic Interview format and coding conventions, have been reported in detail elsewhere (29, 30). The test-retest reliability of the diagnostic modules of the Munich-Composite International Diagnostic Interview was fair to good, with kappa values ranging from 0.64 (Youle’s Y=0.80) to 0.78 (Youle’s Y=0.82). In tests of validity, a good concordance between clinicians’ diagnoses and interview DSM-IV diagnoses was found for all disorders (kappa=0.50–0.96) except psychotic disorders (kappa=0.21).

Posttraumatic stress disorder (and other mental disorders) was defined according to the DSM-IV criteria by using the Munich-Composite International Diagnostic Interview diagnostic algorithm (27). Details of PTSD diagnosis have been presented previously (5). Briefly, a screening question, a written list of 10 groups of potentially traumatic events, and an open-ended question about any other traumatic events (to avoid the necessity of speaking about embarrassing and stigmatizing traumas) were presented, followed by questions about the presence of the DSM-IV A2 criterion (intense fear, helplessness, or horror) during each reported event and questions that probed for the most severe (i.e., most distressing to the individual) event, as well as linkages between events. If the respondent indicated several qualifying events (fulfilling the DSM-IV A1 and A2 criteria) that did not cluster, only the criteria for the most distressing event were assessed by asking which of the events had been the most upsetting at the time it occurred. For this report, the 10 specific event types plus the open category were grouped into a category with two exclusive groups: assaultive events (horrific experience during war, imprisonment, being taken hostage or kidnapped, physical attacks, sexual abuse, and rape) and nonassaultive events (serious accidents, experience of natural catastrophes, sudden death or threat of death of associates, and witnessing traumatic events that happened to others). Similar categories have been used in other studies (31). One-week test-retest reliability of the Munich-Composite International Diagnostic Interview PTSD section was acceptable (kappa=0.79), as was the validity (kappa=0.85) assessed by using the diagnostic concordance between the Munich-Composite International Diagnostic Interview diagnosis and the clinical diagnosis (29, 30).

For this report, we specified two additional categories: “subthreshold PTSD” and “PTSD symptoms.” The category of subthreshold PTSD refers to persons who fulfilled the A, B (traumatic event, fear, and persistent reexperiences), and E criteria (duration) for DSM-IV PTSD but did not completely fulfill the C (avoidance or numbing of general responsiveness) and/or D criteria (increased arousal), although they reported at least one symptom of each of the C and D criteria categories with a duration of more than 1 month. The DSM-IV criterion of impairment was not required (although it was met by some respondents with subthreshold PTSD). Similar subthreshold diagnoses of PTSD have been discussed in the literature (32). The category of “PTSD symptoms” at follow-up refers to persons who fulfilled the A and the B criteria of DSM-IV PTSD but did not meet the criteria for full or subthreshold PTSD. The course specifiers of a chronic course or remission refer to the time between the baseline and second follow-up interviews and describe transitions from subthreshold PTSD to full DSM-IV PTSD, unchanged status between baseline and second follow-up, or transitions to remission (not fulfilling the criteria for the categories defined earlier). Partial remission with PTSD symptoms was included in the category of chronic course. Only persons who failed to meet the criteria for DSM-IV PTSD or our operational criteria for subthreshold PTSD or PTSD symptoms (as described earlier) were considered to have remission for the purpose of this investigation.

Statistical Analysis

Data were weighted to consider different sampling probabilities as well as systematic nonresponse at baseline. The Stata software package (33) was used to calculate proportions and standard errors as well as robust confidence intervals (CIs) for weighted data. Multiple logistic regression analyses with odds ratios were used to describe significant differences between the respondents with remission and those with a chronic course. Logistic regression analyses with adjustment for age and gender were used to calculate odds ratios for incident disorders (i.e., first onset) in the two course specifier groups; the total follow-up sample (without respondents with PTSD) was used as the reference group. Persons with a preexisting case (at baseline) of any of the disorders considered in the analysis of incident disorders were excluded from each of these regression analyses because they were not at risk for the incidence of that disorder. In addition, to avoid confounding findings with new traumatic events, we conducted all logistic regression analyses using data from a sample that excluded persons who experienced new traumatic events between baseline and follow-up. We did not find selective attrition from baseline to follow-up relevant to our categories of DSM-IV PTSD and subthreshold PTSD.

Results

Diagnostic and Symptom Status From Baseline to the Second Follow-Up

As Table 1 shows, at baseline 5.7% (N=125) of the respondents age 14–24 years fulfilled the criteria for either subthreshold PTSD (4.4%) or full DSM-IV PTSD (1.3%). The proportion of female respondents was significantly higher among respondents with full PTSD (83.9%) (odds ratio=5.22, 95% CI=1.61–16.80).

A total of 56.7% of respondents with subthreshold PTSD at baseline and 42.9% of those with full PTSD at baseline reported not meeting the criteria for PTSD symptoms at either the first or the second follow-up assessment, for an
overall remission rate of 52% (N=65). The probability of having full DSM-IV PTSD at the second follow-up was highest among responders with baseline full PTSD (25.6%); only 4.2% of those with subthreshold PTSD at baseline progressed to full PTSD. The remaining 48% of those with full or subthreshold PTSD at baseline reported experiencing full or subthreshold PTSD or PTSD symptoms during the follow-up periods and were classified as having a chronic course (N=60). There were no significant differences in the distribution of responders with subthreshold PTSD and with full PTSD in the two course specifier groups (difference between remission and a chronic course among responders with subthreshold PTSD: odds ratio=0.57, 95% CI=0.21–1.59; difference between remission and a chronic course among responders with full PTSD: odds ratio=1.74, 95% CI=0.62–4.84). However, 44.5% of those with subthreshold PTSD and 22.6% of those with full PTSD experienced new traumatic events during the follow-up period (odds ratio=2.57, 95% CI=1.64–4.02; odds ratio=0.89, 95% CI=0.36–2.22, respectively).

**Differences Between Complete Remission and a Chronic Course**

Table 2 shows differences in certain baseline PTSD characteristics and additional factors between responders with full remission and those with a chronic course. Overall, respondents with a chronic course had higher rates of most of the PTSD characteristics and other risk factors that were assessed, except for assaultive trauma types, number of symptoms, and presence of other baseline mental disorders.

To identify core variables associated with chronicity, we performed multiple logistic regression analyses with all significant variables from Table 2. The experience of new traumatic events between baseline and follow-up was the most robust and significant difference between the two course specifier groups (odds ratio=5.21, 95% CI=1.95–13.92). To avoid confounding between the other predictors and new traumatic events, we also performed the multiple logistic regression analyses using data from the subsample of respondents who had not experienced new traumatic events. Significant differences between the two course specifier groups were found for avoidant symptoms (odds ratio=10.16, 95% CI=1.73–59.51), help seeking (odds ratio=5.50, 95% CI=1.04–29.05), and self-control (odds ratio=2.31, 95% CI=1.04–29.05). The number of symptoms, however, was negatively associated with chronicity (odds ratio=0.15, 95% CI=0.02–0.92).

**Incident Disorders During Follow-Up**

Table 3 shows associations between the two course specifiers and other incident mental disorders with onset during the follow-up period among persons who had never before fulfilled the criteria for these disorders. We did not directly compare differences between the two groups, but we compared rates of onset of new disorders in the two groups with the rates of incident disorders during follow-up in the total sample of respondents with no PTSD and no history of the specific disorder at baseline. In addition, to avoid confounding of new traumatic events and the onset of other disorders, we conducted the analysis with data from a subset of persons who had not experienced new traumatic events.

Table 2 shows that both course specifier groups experienced other incident disorders during the follow-up period, but the overall rate was not significantly different in the category “any disorder” from the rate of follow-up disorders in the reference group. However, although we found no significant odds ratios for specific incident disorders among respondents with remission of PTSD at follow-up, a chronic PTSD course was significantly associated with higher rates of incident somatoform disorders (odds ratio=4.24, 95% CI=1.60–11.19) and incident other anxiety disorders (odds ratio=4.07, 95% CI=1.15–14.37) during the follow-up period. Table 3 also shows

---

### TABLE 3. Incident Other Disorders During 34–50-Month Follow-Up Among Respondents With Posttraumatic Stress Disorder (PTSD) at Baseline and With Complete Remission or a Chronic Course in a Community Sample of Adolescents and Young Adults Age 14–24 Years

<table>
<thead>
<tr>
<th>Incident Disorders</th>
<th>Reference Groupb</th>
<th>Adjusted Odds Ratio 95% CI</th>
<th>Respondents With PTSD Remission at Follow-Up (N=65)</th>
<th>Adjusted Odds Ratio 95% CI</th>
<th>Respondents Without New Traumas (N=50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any disorder</td>
<td>N=625</td>
<td>Weighted % N=10</td>
<td>67.3 % N=3.07 95% CI=0.97–9.69</td>
<td>N=65</td>
<td>2.82 % N=0.62–12.88</td>
</tr>
<tr>
<td>Substance use disorders</td>
<td>N=406</td>
<td>Weighted % N=8</td>
<td>19.6 % N=1.60 95% CI=0.61–4.17</td>
<td>N=50</td>
<td>1.97 % N=0.61–6.39</td>
</tr>
<tr>
<td>Mood disorders</td>
<td>N=292</td>
<td>Weighted % N=4</td>
<td>7.8 % N=0.58 95% CI=0.20–1.70</td>
<td>N=30</td>
<td>0.44 % N=0.10–1.93</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>N=229</td>
<td>Weighted % N=5</td>
<td>23.3 % N=2.66 95% CI=0.86–8.26</td>
<td>N=20</td>
<td>2.57 % N=0.59–11.22</td>
</tr>
<tr>
<td>Somatoform disorders</td>
<td>N=602</td>
<td>Weighted % N=16</td>
<td>29.6 % N=1.39 95% CI=0.68–2.84</td>
<td>N=60</td>
<td>1.47 % N=0.62–3.47</td>
</tr>
</tbody>
</table>

a Includes respondents with both subthreshold PTSD and full DSM-IV PTSD. Respondents with subthreshold PTSD fulfilled the A (traumatic event, fear), B (persistent reexperiences), and E (duration) criteria for DSM-IV PTSD but did not fulfill completely the C (avoidance or numbing of general responsiveness) and/or D (increased arousal) criteria, although they reported at least one symptom in each of the C and D criteria categories with a duration of more than 1 month.
b Includes all respondents in the overall community sample at follow-up who did not have PTSD and did not at baseline fulfill the criteria for the incident disorders under consideration.
that new traumatic events were specifically associated with incident mood disorders among those with a chronic course (odds ratio=2.46, 95% CI=1.12–5.41).

Discussion

The key findings of the study are as follows: More than one-half of the sample with full DSM-IV PTSD at baseline remained symptomatic for more than 3 years, and more than one-half fulfilled the criteria for subthreshold PTSD or DSM-IV PTSD at 34–50-month follow-up. Nearly one-half of those with subthreshold PTSD at baseline remained symptomatic or fulfilled the criteria for subthreshold PTSD or DSM-IV PTSD at follow-up. These estimates are close to those from other epidemiological studies in older populations (1, 14).

Compared to our previous findings and findings from other studies, the predictors of course partly seemed to differ from the predictors of traumatic events and predictors of the onset of PTSD. For example, an early onset and number of traumas were not significantly associated with the course of PTSD in this sample. However, the experience of new traumas during the follow-up interval distinguished a chronic course from a more favorable course with remission. Furthermore, a higher number of avoidant symptoms (from cluster C) at baseline predicted a chronic course. Higher self-competence, which may indicate a greater ability to cope with the effects of recurrent or new traumatic events, was associated with a lower risk of chronicity. It remains to be shown whether these results are stable during a longer course of illness or whether they reflect fluctuating symptoms (14). It is further noteworthy that we found no significant differences for other baseline disorders between respondents with remission and those with a chronic course. This finding might be related to the high rate of comorbidity of other disorders at baseline in both groups (more than 50%).

However, a chronic course of PTSD seems to be significantly associated with incident disorders during follow-up. A higher risk of somatoform disorders or syndromes in persons with PTSD has often been described (34) and was associated with a chronic course in our study. The lack of association between mood disorders and chronicity was somewhat surprising and may be an artifact of the low number of cases of depression in this young adult sample. The role of depression in relation to traumatic events and PTSD has been discussed in the literature, and further empirical evidence in larger samples is required, especially with regard to the course of PTSD (22). Although previous research found associations between substance use disorders and PTSD (5), we could not show that incident substance use disorders were significantly related to course during this relatively brief follow-up period.

Some limitations of the study should be addressed. The findings were restricted to a small number of respondents with full syndromal cases of DSM-IV PTSD at baseline and follow-up. But, as we have noted previously, in this representative community sample of adolescents and young adults, the prevalence of full-blown DSM-IV PTSD is lower than in other epidemiological studies (5). It is also possible that our findings from this relatively young, urban German community sample, which consisted of well-educated persons from an area with a relatively high economic status, may not generalize to other populations.

Our definition of remission refers to the lack of PTSD criterion B (reexperiencing) and E (duration), but it is possible that some persons might have residual symptoms confined to avoidance and/or hyperarousal. Although one might question whether lack of criterion B and E symptoms constitutes remission, we felt that continuing to refer to these persons as having PTSD-related symptoms was less than parsimonious, because such symptoms might be attributable to other DSM disorders. Finally, our findings on course may not generalize to the course after exposure to specific types of traumatic events that were rare or did not occur at all in the present sample (such as natural catastrophes or terrorist attacks).

In conclusion, the results of this longitudinal study confirm in a prospective cohort that PTSD is often a persistent and chronic disorder. In adolescents and young adults, exposure to new traumatic events and seeking help for PTSD symptoms (which may be either an indicator of severity or of coping ability) are associated with poorer outcomes. Avoidant symptoms in particular seem to predict a chronic course. Efforts to prevent persons from being exposed to new traumatic events during the course of PTSD could lessen the chronicity of this disorder. Prevention may be achieved through the implementation of therapies that include techniques for teaching individuals how to seek safe living environments and nonabusive social and romantic relationships (35). The role of other comorbid disorders (e.g., anxiety and somatoform disorders) in influencing the course of PTSD and the possibility that con-
current treatment of these symptom domains might re-
duce chronicity require further investigation.

References

1. Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB: Post-
traumatic stress disorder in the National Comorbidity Survey. Arch Gen Psychiatry 1995; 52:1048–1060
3. Shore JJ, Vollmer WM, Tatum EJ: Community patterns of post-
9. Kilpatrick DG, Saunders BE, Veronen LJ, Best CL, Von JM: Crimi-
nal victimization: lifetime prevalence, reporting to police, and psychological impact. Crime Delinqu 1987; 33:479–489
13. Asarnow J, Glynn S, Pyoons RS, Nahum J, Guthrie D, Cantwell DP, Franklin B: When the earth stops shaking: earthquake se-
quelea among children diagnosed for pre-earthquake psycho-
15. Green BL, Lindy JD, Grace MC, Glessner GC, Leonard AC, Korol M, Winget C: Buffalo Greek survivors in the second decade: stabil-
18. Mayou RA, Ehlers A, Bryant B: Posttraumatic stress disorder af-
19. Widows MR, Jacobson PB, Fields K. Relation of psychological vulnerability factors to posttraumatic stress disorder symptom-
20. Maercker A, Schützwohl M: Long-term effects of political im-
21. Breslau N, Davis GC, Peterson EL, Schultz L: Psychiatric se-
quelea of posttraumatic stress disorder in women. Arch Gen Psychiatry 1997; 54:81–87
22. Breslau N, Davis GC, Peterson EL, Schultz LR: A second look at comorbidity in victims of trauma: the posttraumatic stress dis-
order–major depression connection. Biol Psychiatry 2000; 48:
902–909
513–523
153
25. Wittchen H-U, Perkonigg A, Lachner G, Nelson CB: The Early Devel-
opment Stages of Psychopathology study (EDSP): a methodo-
tuntsersuchung (DIA-X-Lifetime); Ergänzungsheft (DIA-X-lifetime); Interviewheft Querschnittuntersuchung (DIA-X 12 
Monate); Ergänzungsheft (DIA-X 12 Monate); PC-Programm zur 
Durchführung des Interviews (Längs- und Querschnittun-
suchung); Auswertungsprogramm. Frankfurt, Germany, Swets 
& Zeitlinger, 1997
28. Lachner G, Wittchen H-U: Skala zum Vergleich von Kompeten-
zonen. München, Max Planck-Institut, Eigendruck, 1995
International Diagnostic Interview (CIDI) correctly identify

1326

http://ajp.psychiatryonline.org

Am J Psychiatry 162:7, July 2005
National Trends in Hospitalization of Youth With Intentional Self-Inflicted Injuries

Mark Olfson, M.D., M.P.H.
Marc J. Gameroff, Ph.D.
Steven C. Marcus, Ph.D.
Ted Greenberg, M.P.H.
David Shaffer, M.D.

Objective: The authors examined national trends from 1990 to 2000 in the utilization of community hospital inpatient services by young people (5–20 years of age) with intentional self-inflicted injuries.

Method: Discharge abstracts from a nationally representative sample of community hospitals were analyzed, with a focus on youth discharges (N=10,831) with a diagnosis of intentional self-inflicted injury (ICD-9-CM: E950–E959). Census data were used to derive national population-based rates of self-inflicted injuries requiring inpatient treatment. Overall population-based trends in hospitalizations for self-inflicted injury were calculated and stratified by gender and age. Among youths hospitalized with a self-inflicted injury, trends were also calculated for length of stay, inpatient costs, method of injury, and associated mental disorder diagnoses.

Results: The annual hospitalization rate of youths with self-inflicted injuries declined from 49.1 per 100,000 in 1990 to 44.9 per 100,000 in 2000, and the mean length of inpatient stay significantly declined from 3.6 days to 2.7 days. Among the hospitalized patients, there were increases in the rate of cutting (4.3% to 13.2%) and ingestion of acetaminophen (22.1% to 26.9%), antidepressants (10.0% to 14.0%), and opiates (2.3% to 3.3%) as a cause of injury, whereas there were decreases in the ingestion of salicylates (14.9% to 10.2%) and barbiturates (1.5% to 0.7%). There were significant increases in the proportion of subjects with primary mental disorder discharge diagnoses of depressive disorder (29.2% to 46.0%), bipolar disorder (1.3% to 8.2%), and substance use disorder (5.4% to 10.7%) and significant decreases in the rate of adjustment disorders (22.2% to 11.4%) and non-mental disorders (31.9% to 13.6%). After excluding cutting, which may be more closely related to self-mutilation than suicidal self-injury, the annual hospitalization rate of youths with self-inflicted injuries declined from 47.2 per 100,000 in 1990 to 39.4 per 100,000 in 2000.

Conclusions: Over the decade of study, young people admitted to community hospitals with self-inflicted injuries tended to have more severe psychiatric diagnoses and to be treated during shorter inpatient stays. These trends suggest that the role of youth inpatient care has narrowed, becoming focused on those with severe psychiatric disorders.

Considerable controversy surrounds the proper role of short-term inpatient psychiatric treatment for the acute care of young people following an intentional self-injury (1). Whereas some clinical researchers have questioned the necessity and usefulness of admitting acutely self-injurious youth for inpatient psychiatric treatment for other than medical reasons (2), others have developed specific psychiatric indications for admitting suicidal youth to the hospital (3). In practice, emergency room clinicians must grapple with determining whether acutely self-injurious children and adolescents can be safely managed from a medical and psychiatric perspective in treatment settings that are less restrictive and disruptive than hospital care. A wide range of clinical considerations—including medical severity of the attempt, child risk factors, availability of family support and community resources, cost and reimbursement issues, and concerns over adherence with follow-up treatment plans—influence the clinical decision of whether to admit a young person for inpatient care following a self-inflicted injury (4, 5).

Over the last several years, a variety of cost-containment mechanisms have been developed to encourage the substitution of less costly outpatient care for more expensive inpatient services. Between 1987 and 1997, the proportion of total national mental health service expenditures devoted to inpatient psychiatric treatment declined from 40.5% in 1987 to 29.6% in 1997 (6). During this period, many managed care plans developed specific privately held criteria for approving inpatient treatment of suicidal patients (7). For patients admitted to the hospital, managed care utilization management techniques also sought to reduce the length and costs of inpatient care (8, 9). At the same time, the number of families covered by managed behavioral health plans increased (8).

In the current study, we examine national trends from 1990 to 2000 in community hospital admissions of young
people 5–20 years of age with intentional self-inflicted injuries. Using nationally representative data, we describe changes in the rates at which young people were admitted to community hospitals with such injuries and characterize changes in the method of injury, clinical diagnoses received, length of inpatient stay, discharge status, inpatient costs, and primary payer. As a result of changes in the managed care environment, we anticipated that there would be a decline in the rate of youths hospitalized with self-inflicted injuries, a decrease in their length of stay, and an increase in the proportion diagnosed with high-risk conditions, such as mood and substance use disorders, strongly associated with youth suicide.

Method

Data were drawn from the 1990 and 2000 nationwide inpatient samples of the Healthcare Costs and Utilization Project (10). The project is sponsored by the Agency for Healthcare Research and Quality and includes over 100 clinical and nonclinical variables. The Healthcare Costs and Utilization Project consisted of 6,268,515 computerized discharge reports from a geographically diverse sample of 882 community hospitals in 1990 and 7,450,992 discharge reports from a diverse sample of 994 hospitals in 2000. Community hospitals include nonfederal short-term general hospitals and academic medical centers but not specialized psychiatric hospitals. Each year the Healthcare Costs and Utilization Project approximates a 20% stratified sample of U.S. community hospitals. Selection into the sample is based on a stratified probability selection of short-stay nonfederal general hospitals. Weights were constructed on the basis of the reciprocal probability of sampling to approximate national estimates. All percentages in this report are weighted to adjust for the sampling probability.

We limited the analysis to data from youths 5–20 years of age who were admitted to the hospital on an urgent or emergent basis and had a discharge diagnosis for intentional self-inflicted injury (ICD-9-CM: E950.0–E959.9). To estimate population rates of hospitalization for self-inflicted injury, population data were culled from the 1990 and 2000 United States Bureau of the Census (11).

Our first goal was to describe the method of injury, primary mental disorder diagnosis, length of stay, and discharge status of youth by gender and age group in 2000. We then described trends in the hospitalization rate of youths with intentional self-injury between 1990 and 2000. We determined rates of intentional self-injury per 100,000 population, both overall and stratified by age and gender. Substantial missing race/ethnicity data and changes in the federal classification of race/ethnicity categories during the study period prevented a meaningful analysis of these variables.

We then examined trends in the distribution of self-injury method. Injuries were first classified by major category: drug ingestion (E950), hanging/suffocating (E953), firearm (E955), gas asphyxiation (E951, E952), cutting (E956), and a residual group of other types of injury (E954, E957, E958, E959). Psychotrophic drug ingestion was subsequently subclassified on the basis of discharge diagnosis codes: anxiolytics/sedatives (barbiturates, benzodiazepines, and others), antidepressants, antipsychotics, opiates, and other/unspecified psychotropic drugs. Nonpsychotropic ingestion was subclassified as analgesics (acetaminophen, salicylates, and other/unspecified) and other nonpsychotropic substances.

We then examined trends for mean length of stay, discharge status, primary payer, and primary mental disorder diagnoses among youth admissions with intentional self-injuries. Primary mental disorder was defined as the first-listed diagnosis that was a mental disorder (ICD-9-CM: 219–320). Mental disorders were classified into depressive disorders (ICD-9-CM: 296.2, 296.3, 298.0, 300.4, 311), adjustment disorders (309), substance use disorders (291, 292, 303, 304, 305), personality disorders (301), conduct disorder (312, 313.81), psychotic disorders (295, 297–299), bipolar disorder (296.0, 296.1, 296.4–296.9), anxiety disorders (300.0, 300.2, 300.3, 308.3, 309.21, 309.81, 313.0), eating disorders (307.1, 307.5), attention deficit hyperactivity disorder (314), and other mental disorders. A separate category was constructed for discharges with no mental disorder diagnosis.

Because cutting or self-mutilation is rarely associated with completed suicide in young people (12) and tends to be less lethal than other forms of self-injury (13), we also examined overall trends in hospitalization of youths with intentional self-injury excluding cutting (E956). Finally, total inpatient expenditures were calculated for all youths admitted with self-inflicted injuries in 1990 and 2000 and for all admissions excluding self-injury by cutting. The Consumer Price Index for medical care was used to inflate 1990 to 2000 dollars (14).

We used the SUDAAN statistical software package (15) to accommodate the complex sampling design and weights from the Healthcare Costs and Utilization Project when calculating means and corresponding standard errors and to calculate 95% confidence intervals (CIs) for the rate estimates.

Results

Gender and Age Distribution

Of Admissions in 2000

In 2000, several gender differences were evident in the method of injury, primary mental disorder diagnosis, and discharge status of youths admitted with self-inflicted injuries (Table 1). Relative to male subjects, female subjects were significantly more likely to be admitted because of harmful ingesting but less likely to be admitted following self-inflicted injuries due to cutting, hanging/suffocating, firearms, or gas asphyxiation. Female subjects were also significantly more likely to be discharged with a primary mental disorder diagnosis of depressive disorder, adjustment disorder, or an eating disorder but were less likely to be discharged with a diagnosis of substance use disorder, psychotic disorder, or attention deficit disorder. In addition, female subjects were significantly more likely than male subjects to be discharged to home and less likely to die in the hospital or be discharged to an inpatient facility other than a short-term hospital or skilled nursing or intermediate care facility. There was no difference in the number of inpatient days in the community hospital between male (mean=2.9, SE=0.1) and female (mean=2.6, SE=0.1) patients.

The pattern of youth admissions for self-inflicted injuries in 2000 also varied by patient age (Table 2). As compared with younger children (5–14 years of age), the older youth were significantly less likely to be admitted following a self-inflicted hanging/suffocating injury. The older youth were more likely than their younger counterparts to be discharged with a primary mental disorder diagnosis of adjustment disorder or a psychotic disorder but less likely to be discharged with an attention deficit disorder diagnosis or no mental disorder diagnosis. Older youth also tended to have a shorter length of inpatient stay (mean=
Method of injury. 2.83, p=0.0047). In 1990 to 39.4 per 100,000 youth population in 2000 (z=−0.97, p=0.33). The overall rate of community hospital inpatient admissions due to self-injury among youths did not significantly change between 1990 (49 per 100,000 youth population) and 2000 (45 per 100,000 youth population) (Table 3). However, the rate of admissions significantly increased for children 5–9 years of age from 0.4 to 2.1 per 100,000 children (z=−3.96, p<0.0001). After excluding admissions for cutting, the overall rate of inpatient care for youth with intentional self-injuries significantly decreased from 47.2 per 100,000 youth population in 1990 to 39.4 per 100,000 youth population in 2000 (z=2.83, p=0.0047).

Method of injury. There was a significant decrease in the proportion of hospitalizations involving drug ingestion over the 10-year period. Significant decreases were specifically observed in the proportion of hospitalizations associated with barbiturates, salicylates, and unspecified nonpsychotropic drugs. During the same period, significant increases were evident in the proportion of hospitalizations associated with ingestions of antidepressant medications, opiates, and acetaminophen (Table 4). There was also a significant increase in the proportion of hospitalizations that involved cutting and hanging/suffocation.

Mental disorder diagnosis. In both study years, mental disorders were the primary discharge diagnosis for most youths hospitalized because of self-injury. Significant increases were specifically observed in the proportion of discharges with depressive disorder, substance use disorder, bipolar disorder, and attention deficit disorder (Table 5). There was also a significant increase in the proportion of discharges with a personality disorder listed as a secondary diagnosis.

Clinical service characteristics. In both 1990 and 2000, the majority of youths admitted because of intentional self-inflicted injuries were discharged to home (Table 5). Although transfer to short-term hospitals became significantly less common, transfers to other inpatient facilities became more common. Comparatively few young people died in the hospital or left the hospital against medical advice in either year. Also in both years, private insurance was the primary payer in a majority of the hospitalizations. Medicaid, which increased as a percentage of youth hospitalizations with intentional self-injury, was the second most common primary payer. The mean length of stay for the hospitalizations significantly declined from 3.6 days (SE=0.2) in 1990 to 2.7 days (SE=0.1) in 2000 (t=15.8, df=10,829, p<0.0001).

Inpatient expenditures. Total estimated inpatient costs in inflation-adjusted 2000 dollars for youth admissions with intentional self-injuries were $167.5 million in 1990 (95% CI=147.3–187.8) and $168.2 million (95% CI=147.2–187.8) in 2000. After excluding admissions for self-injury due to cutting, the respective estimates were $160.8 million (95% CI=141.4–180.2) in 1990 and $149.7 million (95% CI=130.3–169.1) in 2000 in inflation-adjusted 2000 dollars.

Discussion

Between 1990 and 2000, there was a statistically nonsignificant decrease in the annual rate of community hospital inpatient service utilization among youths admitted following intentional self-injury (from approximately 49 per 100,000 in 1990 to 45 per 100,000). This trend roughly parallels the national decline in suicides among youths 15–19 years of age, from 11.1 per 100,000 in 1990 to 8.2 per 100,000 in 2000 (12).

The trend in hospital admissions may portray important changes in the care of young people who intentionally injure themselves. During the period from 1991 to 2001, the Youth Risk Behavior Survey reported a substantial increase in the rate of injurious suicide attempts by students.

### TABLE 1. Clinical Characteristics of Youths With Self-Inflicted Injuries Admitted to Community Hospital Inpatient Services in 2000, by Patient Gender

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Rate (%)</th>
<th>Rate (%)</th>
<th>χ²</th>
<th>df=1</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method of injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingestion</td>
<td>71.3</td>
<td>87.3</td>
<td>81.7</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cutting</td>
<td>17.1</td>
<td>11.5</td>
<td>20.8</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hanging/suffocating</td>
<td>3.5</td>
<td>0.3</td>
<td>34.6</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Firearm</td>
<td>2.7</td>
<td>0.2</td>
<td>25.4</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Gas asphyxiation</td>
<td>1.3</td>
<td>0.1</td>
<td>17.3</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Other</td>
<td>5.1</td>
<td>1.9</td>
<td>24.8</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Primary mental disorder diagnosis at discharge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive disorder</td>
<td>39.2</td>
<td>49.0</td>
<td>31.5</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Adjustment disorder</td>
<td>8.6</td>
<td>12.6</td>
<td>19.4</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Substance use disorder</td>
<td>16.2</td>
<td>8.3</td>
<td>55.6</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Personality disorder</td>
<td>0.8</td>
<td>0.9</td>
<td>0.03</td>
<td>0.86</td>
<td>0.67</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>1.7</td>
<td>1.5</td>
<td>0.18</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>Psychotic disorder</td>
<td>3.7</td>
<td>1.1</td>
<td>23.2</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>9.9</td>
<td>7.4</td>
<td>6.8</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>1.0</td>
<td>1.3</td>
<td>1.1</td>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td>Eating disorder</td>
<td>0.0</td>
<td>0.5</td>
<td>23.9</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Attention deficit disorder</td>
<td>2.3</td>
<td>0.6</td>
<td>16.5</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Other mental diagnosis</td>
<td>3.0</td>
<td>3.1</td>
<td>0.04</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>No mental diagnosis</td>
<td>13.7</td>
<td>13.6</td>
<td>0.02</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>Discharge status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home</td>
<td>62.5</td>
<td>69.1</td>
<td>18.8</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Short-term hospital</td>
<td>6.4</td>
<td>6.9</td>
<td>0.5</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>Skilled nursing/intermediate facility</td>
<td>0.9</td>
<td>0.4</td>
<td>3.6</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Other inpatient facility</td>
<td>26.5</td>
<td>22.1</td>
<td>12.8</td>
<td>0.0004</td>
<td>0.0004</td>
</tr>
<tr>
<td>Against medical advice</td>
<td>1.9</td>
<td>1.3</td>
<td>2.1</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Died in hospital</td>
<td>1.7</td>
<td>0.2</td>
<td>21.9</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Youth defined as 5–20 years of age. Rates and means are nationally weighted estimates from the Healthcare Costs and Utilization Project national inpatient sample for 2000. SUDAAN software was used to account for the complex survey design.
This trend should alert clinicians to the risks of suicidal behavior in young children. While the cause of this increase remains unknown, some evidence links child exposure to video games and movie violence to violent attitudes (24) and behavior (25). One small study has suggested that younger children who attempt suicide report many of the same depressive symptoms common to suicidal adolescents (26). The recent increase of inpatient admissions of young children with intentional self-inflicted injuries highlights the importance of prevention and early intervention programs that target preschool (27) or early grade school (28, 29) children at risk for mental health problems.

The changing diagnostic profile of young people admitted to the hospital following intentional self-injury could indicate a more focused approach for especially high-risk youth. Between 1990 and 2000, the proportion of discharges in which a mood disorder or substance use disorder was the first listed mental disorder significantly increased, whereas the proportion of adjustment disorders or no mental disorder discharges significantly declined. Psychological autopsy studies reveal that mood disorders occur in approximately two-thirds of youth suicides (30–32) and that substance use disorders occur in up to two-thirds of older boys who complete suicide (30, 32). By contrast, young people with adjustment disorders or no mental disorder discharges are considerably less common among youth suicides (30, 32).

A trend toward more severe mental disorder diagnoses among inpatient youth discharges suggests the importance of developing rapid and efficient diagnostic procedures to identify young people with high-risk conditions (33). In one recent study, a school-based program for high school students that focused on depression and suicide risk was associated with a reduction in suicide attempts (34).

### Table 2. Clinical Characteristics of Youths With Self-Inflicted Injuries Admitted to Community Hospital Inpatient Services in 2000, by Patient Age

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Rate (%)</th>
<th>Rate (%)</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Among Patients</td>
<td>Among Patients</td>
<td>(df=1)</td>
</tr>
<tr>
<td>Method of injury</td>
<td></td>
<td></td>
<td>p</td>
</tr>
<tr>
<td>Ingestion</td>
<td>81.1</td>
<td>82.8</td>
<td>1.2</td>
</tr>
<tr>
<td>Cutting</td>
<td>13.9</td>
<td>13.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Hanging/suffocating</td>
<td>2.4</td>
<td>1.0</td>
<td>6.9</td>
</tr>
<tr>
<td>Firearm</td>
<td>0.7</td>
<td>1.1</td>
<td>1.4</td>
</tr>
<tr>
<td>Gas asphyxiation</td>
<td>0.4</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Other</td>
<td>3.0</td>
<td>2.9</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### Table 3. Rate of Community Hospital Inpatient Service Utilization by Youths With Self-Inflicted Injuries in 1990 and 2000, by Age and Gender

<table>
<thead>
<tr>
<th>Groupb</th>
<th>Hospitalization Rate (per 100,000 youth population)</th>
<th>1990</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (N=10,831)</td>
<td>49.1</td>
<td>44.9</td>
<td>53.3</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (N=7,304)</td>
<td>29.6</td>
<td>26.7</td>
<td>32.6</td>
</tr>
<tr>
<td>Female (N=7,527)</td>
<td>69.6</td>
<td>63.1</td>
<td>76.1</td>
</tr>
</tbody>
</table>

†Youth defined as 5–20 years of age. Inpatient treatment is limited to admission type urgent and emergent. Weighted rates were constructed as follows: for each cell, the numerator was a nationally weighted estimate from the Healthcare Costs and Utilization Project national inpatient sample from 1990 or 2000. The denominator for each cell was a U.S. Census population figure from 1990 or 2000. Confidence intervals took into account the standard error of the Healthcare Costs and Utilization Project estimate; the U.S. Census data were from a 100% enumeration and do not have standard errors associated with them. SUDAAN software was used to account for the complex survey design.

Ns are unweighted.
Diagnosis. Bipolar disorder in young people has been associated with inappropriate utilization management policies. The diagnostic trends are simply a coding response to reimbursement changes, as indicated by the observed increases in the proportion of youth admissions with a primary mental disorder diagnosis. For example, among youth admissions with a primary mental disorder diagnosis, the shortening of inpatient treatment may be related to the trend toward limiting care to those with more severe illness (37, 38), and concern exists that there has been a trend toward overdiagnosis of youth bipolar disorder (39). From the available data, it is not possible to determine the extent to which the increase in discharges associated with bipolar disorder represents a true change in diagnostic composition as opposed to a change in diagnostic practices. Little attention has thus far been focused on early intervention in young people with bipolar disorder (40).

There was also a trend toward shorter hospitalizations for young people with self-inflicted injuries. Together with the trend toward limiting care to those with more severe diagnoses, the shortening of inpatient treatment may be placing inpatient staff under increased time pressures to locate appropriate outpatient care. Under these constraints, it is perhaps not surprising that an increasing proportion of inpatients were transferred to other inpatient facilities.

The trends in mental disorder diagnoses among youths hospitalized following self-inflicted injuries may reflect broad changes in diagnostic practices or clinical decision making in response to managed care restrictions on inpatient care for less severe mental disorders. To explore this possibility, we performed a set of post hoc Healthcare Costs and Utilization Project analyses. The trends observed among admissions with self-inflicted injuries were also apparent in the larger sample of youth admissions with a primary mental disorder diagnosis. For example, among youth admissions with primary mental disorder diagnoses, the proportion who were diagnosed with an adjustment disorder declined from 1990 (16.8%) to 2000 (6.5%), whereas increases were seen in depressive disorder (25.7% to 34.4%) and bipolar disorder (4.2% to 12.2%). Because similar trends were further observed among youth whose admissions were self-pay or not charged (adjustment disorder: 19.4% to 7.9%; depressive disorder: 18.4% to 31.1%; bipolar disorder: 2.0% to 9.6%), it is unlikely that the diagnostic trends are simply a coding response to restrictive utilization management policies.

During the decade under study, there was a particularly impressive increase in the proportion of hospitalizations in which bipolar disorder was the leading mental disorder diagnosis. Bipolar disorder in young people has been associated with an increased risk of suicide attempts (35) and completion (2, 36). In one case/control study of youth suicide, the odds ratio of suicide completion for bipolar disorder approached that for substance use disorders (36). However, manic or manic-like symptoms in young people may be difficult to distinguish from symptoms of ADHD (37, 38), and concern exists that there has been a trend toward overdiagnosis of youth bipolar disorder (39). From the available data, it is not possible to determine the extent to which the increase in discharges associated with bipolar disorder represents a true change in diagnostic composition as opposed to a change in diagnostic practices. Little attention has thus far been focused on early intervention in young people with bipolar disorder (40).

### Table: Hospitalization Trends for Self-Injuring Youth

<table>
<thead>
<tr>
<th>Method of Injury</th>
<th>Rate (%)</th>
<th>Rate (%)</th>
<th>p</th>
<th>χ² (df=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug ingestion</td>
<td>92.8</td>
<td>82.4</td>
<td>0.45</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Psychotropic drug ingestion</td>
<td>31.1</td>
<td>33.6</td>
<td>4.2</td>
<td>0.04</td>
</tr>
<tr>
<td>Anxiety/sedatives</td>
<td>9.5</td>
<td>9.2</td>
<td>0.2</td>
<td>0.70</td>
</tr>
<tr>
<td>Barbiturates</td>
<td>1.5</td>
<td>0.7</td>
<td>10.3</td>
<td>0.001</td>
</tr>
<tr>
<td>Benzodiazepines</td>
<td>5.5</td>
<td>6.2</td>
<td>1.4</td>
<td>0.24</td>
</tr>
<tr>
<td>Other/unspecified</td>
<td>2.5</td>
<td>2.5</td>
<td>0.0</td>
<td>0.83</td>
</tr>
<tr>
<td>Antidepressants</td>
<td>10.0</td>
<td>14.0</td>
<td>31.8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Antipsychotics</td>
<td>2.4</td>
<td>2.5</td>
<td>0.2</td>
<td>0.66</td>
</tr>
<tr>
<td>Opiates</td>
<td>2.3</td>
<td>3.3</td>
<td>8.0</td>
<td>0.005</td>
</tr>
<tr>
<td>Other/unspecified</td>
<td>10.6</td>
<td>12.0</td>
<td>2.8</td>
<td>0.10</td>
</tr>
<tr>
<td>Nonpsychotropic drug ingestion</td>
<td>69.9</td>
<td>60.2</td>
<td>41.7</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Alcohol</td>
<td>44.0</td>
<td>45.4</td>
<td>1.2</td>
<td>0.28</td>
</tr>
<tr>
<td>Acetaminophen</td>
<td>22.1</td>
<td>26.9</td>
<td>16.9</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Salicylates</td>
<td>14.9</td>
<td>10.2</td>
<td>28.6</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Other/unspecified</td>
<td>9.6</td>
<td>11.1</td>
<td>4.6</td>
<td>0.03</td>
</tr>
<tr>
<td>Other substances</td>
<td>36.5</td>
<td>28.4</td>
<td>38.8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cutting</td>
<td>4.3</td>
<td>13.2</td>
<td>39.8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hanging/suffocating</td>
<td>0.6</td>
<td>1.3</td>
<td>10.1</td>
<td>0.002</td>
</tr>
<tr>
<td>Firearm</td>
<td>0.9</td>
<td>1.0</td>
<td>0.1</td>
<td>0.71</td>
</tr>
<tr>
<td>Gas asphyxiation</td>
<td>0.6</td>
<td>0.5</td>
<td>0.1</td>
<td>0.76</td>
</tr>
<tr>
<td>Other</td>
<td>1.3</td>
<td>2.9</td>
<td>28.2</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

### Table: Community Hospital Inpatient Service Utilization Characteristics in 1990 and 2000 for Youths With Self-Inflicted Injuries

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Rate (%)</th>
<th>Rate (%)</th>
<th>p</th>
<th>χ² (df=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary mental diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive disorder</td>
<td>29.2</td>
<td>46.0</td>
<td>64.1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Adjustment disorder</td>
<td>22.2</td>
<td>11.4</td>
<td>59.7</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Substance use disorder</td>
<td>5.4</td>
<td>10.7</td>
<td>71.1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Personality disorder</td>
<td>2.0</td>
<td>0.8</td>
<td>8.9</td>
<td>0.003</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>1.6</td>
<td>1.6</td>
<td>0.0</td>
<td>0.97</td>
</tr>
<tr>
<td>Psychotic disorder</td>
<td>1.3</td>
<td>1.9</td>
<td>3.1</td>
<td>0.08</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>1.3</td>
<td>8.2</td>
<td>97.4</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>1.1</td>
<td>1.2</td>
<td>0.1</td>
<td>0.80</td>
</tr>
<tr>
<td>Eating disorder</td>
<td>0.3</td>
<td>0.4</td>
<td>0.2</td>
<td>0.62</td>
</tr>
<tr>
<td>Attention deficit disorder</td>
<td>0.1</td>
<td>1.1</td>
<td>36.9</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Other mental diagnosis</td>
<td>3.6</td>
<td>3.1</td>
<td>1.0</td>
<td>0.33</td>
</tr>
<tr>
<td>No mental diagnosis</td>
<td>31.9</td>
<td>13.6</td>
<td>107.3</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Personality disorder (not necessarily primary)</td>
<td>6.2</td>
<td>8.4</td>
<td>7.7</td>
<td>0.006</td>
</tr>
<tr>
<td>Discharge status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home</td>
<td>69.7</td>
<td>67.1</td>
<td>2.0</td>
<td>0.15</td>
</tr>
<tr>
<td>Short-term hospital</td>
<td>9.5</td>
<td>6.8</td>
<td>8.5</td>
<td>0.004</td>
</tr>
<tr>
<td>Skilled nursing/intermediate facility</td>
<td>1.2</td>
<td>0.6</td>
<td>3.9</td>
<td>0.05</td>
</tr>
<tr>
<td>Other inpatient facility</td>
<td>15.9</td>
<td>23.4</td>
<td>22.3</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Against medical advice</td>
<td>3.2</td>
<td>1.5</td>
<td>14.3</td>
<td>0.0002</td>
</tr>
<tr>
<td>Died in hospital</td>
<td>0.4</td>
<td>0.6</td>
<td>1.6</td>
<td>0.21</td>
</tr>
<tr>
<td>Primary payer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private insurance</td>
<td>51.1</td>
<td>59.5</td>
<td>17.2</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Medicaid</td>
<td>18.8</td>
<td>22.3</td>
<td>5.6</td>
<td>0.02</td>
</tr>
<tr>
<td>Self-pay or no charge</td>
<td>9.2</td>
<td>13.8</td>
<td>18.7</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Medicare</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.59</td>
</tr>
<tr>
<td>Other</td>
<td>20.4</td>
<td>4.0</td>
<td>98.4</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

*Youth defined as 5–20 years of age. Rates and means are nationally weighted estimates from the Healthcare Costs and Utilization Project national inpatient sample for 1990 and 2000. Numbers in parentheses are unweighted Ns. There were a number of missing observations in 1990 and 2000 for discharge status (N=13 and 130, respectively) and primary payer (N=33 for both years). SUDAAN software was used to account for the complex survey design.*
In the acute outpatient management of suicidal young people, encouraging results have been reported with rapid-response outpatient psychiatric teams (41), home-based interventions (42, 43), interpersonal problem-solving skills training (44), and brief cognitive behavior therapy for family members (45). It is not known whether access to these and other relevant outpatient mental health services has expanded to help compensate for the narrowing role of short-term hospitalization in the care of young people with intentional self-inflicted injuries.

Several changes occurred in the pattern of intentional self-injury methods. Methods associated with high case fatality rates, including gas asphyxiation, hanging/suffocating, and firearms (23), remained relatively uncommon, possibly because of deaths in the community or the emergency room before hospital admission. Although ingestions declined as a proportion of admissions for intentional self-injury, there were proportionate increases in ingestion of acetaminophen, antidepressants, and opiates. These proportionate increases, together with proportionate decreases in barbiturate and salicylate ingestions, may reflect changes in the general use of these substances (46). The increase in acetaminophen ingestion is especially noteworthy because it poses a serious risk of potentially fatal hepatic toxicity (47) that may not be appreciated by young people (48).

During the study period, there was also an increase in the inpatient treatment of intentional self-inflicted injury involving cutting. Self-mutilation in young people tends to have a very low potential for lethality (23). In the United States, six adolescents (age range: 15–19 years) committed suicide by injury with a sharp object in 2000 (12). As compared with other types of self-injury, self-mutilation is clinically associated with greater patient perceived likelihood of rescue (13), lower perceived certainty of death (13), and lower rates of mood disorders (49). When these presumably less clinically severe discharges were excluded from the analysis, there was a statistically significant decline in the rate of youths hospitalized with intentional self-injury: from 47.2 to 39.4 discharges per 100,000 youth.

This study has several limitations. First, diagnostic data are only a crude index of illness severity and risk of subsequent suicide. No information was available concerning several known suicide risk factors in young people including prior suicide attempts (50), precipitating stressful life events (51, 52), access to firearms (2), and family psychiatric history (51, 53). Second, no independent assessment was available of the clinical diagnostic codes. A growth in the inpatient treatment of intentional self-inflicted injuries. However, the proportion of young self-injurious inpatients with high-risk conditions (including depressive, bipolar, and substance use disorders) has increased, and the length of their inpatient stays has declined. These changes, especially in light of increasing injurious youth suicide attempts in the community, indicate that inpatient care may have assumed a narrower and more limited role in the treatment of suicidal young people. As mental health care professionals have come to rely less extensively on inpatient treatment for the acute management of self-injurious young people, community service needs have likely increased for rapid-response, crisis-oriented outpatient care.

References
HOSPITALIZATION TRENDS FOR SELF-INJURING YOUTH

Racial and Ethnic Differences in Utilization of Mental Health Services Among High-Risk Youths

Ann F. Garland, Ph.D.
Anna S. Lau, Ph.D.
May Yeh, Ph.D.
Kristen M. McCabe, Ph.D.
Richard L. Hough, Ph.D.
John A. Landsverk, Ph.D.

Objective: Racial and ethnic disparities in mental health service use have been identified as a major public health problem. However, the extent to which these disparities may be accounted for by other confounding sociodemographic or clinical predictors of service use (e.g., family income, functional impairment, caregiver strain) is relatively unexplored, especially for youth services. The goal of this study was to test for racial/ethnic disparities in use of a variety of outpatient, inpatient, and informal mental health services among high-risk youths, with the effects of other predictive factors controlled.

Method: Participants were 1,256 youths ages 6–18 years who received services in a large, publicly funded system of care (including the child welfare, juvenile justice, special education, alcohol and drug abuse, and mental health service sectors). Youths and caregivers were interviewed with established measures of mental health service use, psychiatric diagnoses, functional impairment, caregiver strain, and parental depression.

Results: Significant racial/ethnic group differences in likelihood of receiving any mental health service and, specifically, formal outpatient services were found after the effects of potentially confounding variables were controlled. Race/ethnicity did not exert a significant effect on the use of informal or 24-hour-care services.

Conclusions: Racial/ethnic disparities in service use remain a public health problem.

Children and adolescents significantly underutilize mental health services, and unmet need for services appears to be greatest among racial/ethnic minority groups (1–3). The Surgeon General’s 2001 report on mental health identified racial/ethnic disparities in mental health service use as a major public health problem (4). However, not all studies have found racial/ethnic discrepancies in mental health service use among youths (5–7). This inconsistency may be related to geographic and/or methodological differences, such as varying attention to potentially confounding predictors or variation in the definition of mental health services across studies.

Analyses of nationally representative survey data revealed lower rates of mental health service use for African American and Latino American children, compared to non-Hispanic white children; follow-up analyses accounting for potentially confounding variables (e.g., insurance status) indicated that unmet need was greatest among Latino American children, compared to non-Hispanic white children (2). However, studies in urban areas with families of relatively homogeneous socioeconomic status found no racial/ethnic differences in service use (6, 7). Likewise, data from a study in a largely rural area revealed no difference in specialty or general service use between African American and non-Hispanic white children (5).

All of these studies included community samples of children and families, as opposed to selected high-risk youths. Very little is known about mental health service use rates for high-risk youths in public service sectors such as child welfare, juvenile justice, and special education services. Youths in these service sectors exhibit high rates of psychiatric disorders and are at high risk for a variety of other maladaptive outcomes (8–10); racial/ethnic minority youths tend to be overrepresented in some of these service sectors (3, 10). Given that they have already been identified in a service sector, these youths are likely to have substantially higher rates of mental health service use, compared to community samples, but the extent to which there is differential utilization by race/ethnicity for different types of services is not known. Service use disparities among high-risk youths would be particularly concerning, given the poor prognosis for youths with untreated psychopathology (11).

Multiple youth-, family-, and system-level factors have been shown to predict youth mental health service use, and many of these factors are associated with race/ethnicity. For example, one of the most consistent predictors of child mental health service use is the caregiver’s perception of the strain of caring for the child (12, 13). Recent evidence has identified significant and robust racial/ethnic differences in caregiver strain (14). Contact with the justice system may be a negative predictor of service use (5, 15), and certain racial/ethnic minority groups are more likely to make contact with the juvenile justice system.
In addition, socioeconomic status and insurance coverage are strongly associated with race/ethnicity and are predictors of service use (2, 17). Therefore, among high-risk youth involved in public service sectors, it is important to examine whether racial/ethnic differences in mental health service use are actually attributable to these potentially confounding variables and/or to clinical factors such as psychiatric diagnoses or functional impairment, which, as expected, are associated with the likelihood of service use (2, 5, 7, 18).

Most studies of racial/ethnic differences in service use have included assessment of some of these potentially confounding variables, but, to our knowledge, this study is the first to make a comprehensive assessment of these variables in a rigorous test of the robustness of racial/ethnic differences, to examine use of a variety of formal and informal types of mental health service use, and to have adequate representation across four major racial/ethnic groups (non-Hispanic white, African American, Latino American, Asian American/Pacific Islander). The purpose of this study was to 1) test for racial/ethnic group differences in mental health service use among high-risk youths in a public system of care and 2) to test whether identified differences persisted when the effects of other factors known to be associated with service use were taken into account.

Method

Participants

Participants were a subsample of 1,256 of the 1,715 youths in the Patterns of Youth Mental Health Care in Public Service Systems study (9). The study was approved by the human subjects protection committee at Children’s Hospital and Health Center in San Diego, San Diego State University, and the University of California, San Diego. The study participants were randomly selected from an enumeration of all youths “active” in one or more of five San Diego County public sectors of care (alcohol and drug abuse, child welfare, juvenile justice, mental health, and public school special education services) for youths with serious emotional disturbance during the first half of 1997 (total population=12,662). The sample was selected by simple random sampling techniques and stratified by race/ethnicity and restrictiveness of care (care in aggregate setting versus the home residence). Data were obtained for 67% of the eligible sample in interviews completed between late 1997 and early 1999. The participants did not differ significantly from the nonparticipants in age, gender, sector affiliation, or racial/ethnic distribution, except that slightly fewer Asian American/Pacific Islanders participated, relative to the eligible sample.

The 1,256 participants who provided data for the analyses reported here included those with complete diagnostic and service use data from adult and child interviews and those who were in one of the following four largest racial/ethnic groups: non-Hispanic whites (N=554, 44%), Latino Americans (N=332, 26%), African Americans (N=282, 22%), and Asian American/Pacific Islanders (N=88, 7%). Race/ethnicity was self-identified for youth ages 11 years and older and was parent-identified for children ages 6–10 years. Eight percent of the total sample of 1,715 (N=135) were missing data because the interviewee was not English speaking (5%, N=88) or because of other reasons (4%, N=68), and 8% (N=145) were missing child interview data because of inability to contact the child or language differences/difficulties.

Two-thirds of the participants were male. The mean age was 13.7 years (SD=3.3). Most of the parent/caregiver informants (hereafter labeled “parents”) were biological parents (72%). Others included adoptive or foster parents, stepparents, and a small number of professional caregivers.

Procedure and Measures

After complete description of the study, written informed consent was obtained from the parent and assent was obtained from the youths. Parents and youths were interviewed individually (usually in their home) about the youth’s mental health use, needs, and a variety of factors potentially associated with mental health service use (e.g., caregiver strain, family income). Parents and youths were compensated (up to $40) for their time, which averaged 3 hours. Interviewer training and reliability checks have been described previously (8, 9).

The measures used in this study were as follows:

Services Assessment for Children and Adolescents. Parent and youth versions of the Services Assessment for Children and Adolescents (19) were used to assess utilization of different types of mental health and substance abuse services. Only past-year service use was examined in this study. A previous analysis showed that the test-retest reliability of the Services Assessment for Children and Adolescents for past-year service use is excellent for parent informants and is good for youth informants age ≥10 years (19). In the current study, service use was considered present if it was reported by either the parent or the youth. The following types of services were assessed:

1. Outpatient services, including specialty outpatient care (e.g., specialty mental health clinics or private providers), nonspecialty outpatient care (e.g., visit to a pediatrician for emotional/behavioral issues), and outpatient alcohol and drug abuse treatment.
2. Twenty-four-hour-care services, including inpatient care in a psychiatric hospital or psychiatric unit within a hospital, residential treatment center/group home services, and inpatient alcohol and drug abuse treatment.
3. Informal services, including self-help groups/peer counseling, counseling from clergy, and services of alternative healers.

National Institute of Mental Health Diagnostic Interview Schedule for Children Version IV. The computer-assisted parent and youth versions of the National Institute of Mental Health Diagnostic Interview Schedule for Children Version IV (20) were used to assess the presence of past-year DSM-IV psychiatric disorders. The reliability and validity of the Diagnostic Interview Schedule for Children are well supported (20). To reduce interview duration, the mood and anxiety modules were administered only to the youths, because youths are likely the best informants for internalizing disorders (21). The disruptive behavior disorder module was administered to both parents and youths, and diagnoses were considered present if either respondent’s report met the diagnostic criteria detailed in the Diagnostic Interview Schedule for Children scoring algorithms (including diagnosis-specific functional impairment).

Composite International Diagnostic Interview substance abuse module. The Composite International Diagnostic Interview substance abuse module (22) was administered to youths age 11 years and older. It was used to identify past-year DSM-IV substance use disorders, including abuse and dependence diagnoses for alcohol, marijuana, hallucinogens, and stimulants. Children younger than age 11 years were assumed to have no substance use.
disorder. The reliability of the Composite International Diagnostic Interview substance abuse module is strong (22).

**Children’s Global Assessment Scale.** The Children’s Global Assessment Scale (23) was used to assess global functional impairment. Interviewers assigned ratings on the Children’s Global Assessment Scale after completion of the youth and parent interviews. The standard cutoff of 60 was used to designate clinically significant functional impairment (24).

**Columbia Impairment Scale.** The Columbia Impairment Scale (25) was used to assess parent- and youth-reported functional impairment. The scale has strong psychometric characteristics, and the recommended clinically significant cutoff score of 15 was used to identify youths with significant functional impairment (25). In this study, youths were considered functionally impaired if they met the cutoff criteria for either the Children’s Global Assessment Scale or the Columbia Impairment Scale.

**Caregiver Strain Questionnaire.** The Caregiver Strain Questionnaire (13) was used to assess the parents’ perceptions of the burden or impact of caring for a child with behavioral problems. The reliability and validity of this 21-item self-report measure are well supported (13). Because there is no recommended clinically significant cutoff score, scores were dichotomized for the analyses by splitting at the median.

**Center for Epidemiologic Studies Depression Scale.** The Center for Epidemiologic Studies Depression Scale (26) was used to assess caregiver depressive symptoms. The 20-item self-report scale has demonstrated strong reliability (coefficient alpha = 0.85–0.90) and validity for use with diverse populations.

**Police contact.** Parents were asked whether the youth had ever been arrested, picked up by police, or given a warning by police. This variable was dichotomized to reflect any versus no lifetime police contact.

**Parental education.** Parents reported on their highest level of education. This variable was also dichotomized to represent those with and without any college-level education.

**Family income.** Total family income was reported by parents on an incremental scale of annual incomes from ≤$1,000 to ≥$200,000. For these analyses, the distribution was divided into quartiles, as follows: first quartile, <$13,000; second quartile, ≤$25,000; third quartile, ≤$45,000; and fourth quartile, >$45,000.

**Insurance status.** Parents reported on their insurance coverage for mental health care. Coverage was categorized as follows: private insurance, government insurance coverage (e.g., MediCal), or no insurance coverage.

### Data Analysis

All analyses were conducted by using Stata (27), with data weighted to represent the enumerated system-of-care population. Chi-square analyses were used initially to test for overall racial/ethnic group differences in rates of service utilization by service category and individual service type. To account for the survey design, the chi-square statistic was converted to an F statistic with noninteger degrees of freedom by using a second-order Rao and Scott’s correction (27, 28). Follow-up logistic regression analyses were used to test for racial/ethnic differences for each service category after the effects of potential confounding variables were controlled. The significance level was set at an alpha of 0.05, and all tests were two-sided.

### Results

The youths in this study had high rates of mental health service utilization. Overall, 72% (N=904) had utilized some type of mental health service within the past year. The most commonly used services were professional outpatient services (64% of the total sample, N=803). Twenty percent (N=251) of the participants had used informal services such as self-help groups, and 13% (N=163) had used inpatient or residential treatment services. Figure 1 displays the rates of any reported mental health service use, as well as rates for outpatient, 24-hour-care, and informal services, for each racial/ethnic group. There were significant differences across racial/ethnic groups for use of any mental health service (χ²=37.1, df=3, p<0.001), outpatient services (χ²=43.0, df=3, p<0.001), and informal services (χ²=20.1, df=3, p<0.001), but not for 24-hour-care services (χ²=7.3, df=3, p=0.23). Non-Hispanic whites had the highest rates of service use for any mental health service and for outpatient services; Asian American/Pacific Islanders had the lowest utilization rates for these categories of service. For informal services use, Latino Americans, Asian American/Pacific Islanders, and non-Hispanic whites had relatively similar rates, and African Americans had the lowest rate.

Table 1 lists the utilization rates by racial/ethnic group for specific types of services within the broad categories reported earlier. Significant differences across racial/ethnic groups were found for specialty mental health outpatient services (χ²=60.6, df=3, p<0.001), outpatient alcohol and drug abuse treatment (χ²=22.1, df=3, p<0.001), inpatient psychiatric hospital treatment (χ²=10.1, df=3, p<0.02), and use of self-help groups (χ²=23.1, df=3, p<0.002). There were no significant racial/ethnic group differences for other nonspecialty outpatient services (e.g., pediatrics/primary care, emergency room, etc.), residential treatment or group
were controlled. After the effects of other potentially confounding variables to the likelihood of use of 24-hour-care or informal services for non-Hispanic whites. Race/ethnicity did not contribute approximately one-half as likely for African Americans (odds ratio=0.54, 95% CI=0.36–0.80) and Asian American/Pacific Islanders (odds ratio=0.48, 95% CI=0.27–0.85) than for non-Hispanic whites. Race/ethnicity did not contribute to the likelihood of use of 24-hour-care or informal services after the effects of other potentially confounding variables were controlled.

Of secondary interest, many of the potentially confounding variables were associated with likelihood of service use. For example, use of any mental health service was positively associated with female gender, higher caregiver strain, contact with alcohol and drug abuse services or mental health services in the year before the study, a DSM-IV non-substance-use diagnosis, and functional impairment. Use of any mental health service was less likely for youths with families in the middle-range incomes, compared to the highest range.

Table 2 presents the results of the regression analyses. The racial/ethnic group variables added significantly to the prediction of likelihood of use of outpatient services and any mental health service, after the effects of all other potentially confounding variables were taken into account. Specifically, the last column indicates that African Americans and Asian American/Pacific Islanders were approximately one-half as likely to receive any mental health service, compared to non-Hispanic whites (African Americans: odds ratio=0.56, 95% confidence interval [CI]=0.37–0.84; Asian American/Pacific Islanders: odds ratio=0.44, 95% CI=0.25–0.77). Outpatient service use was also approximately one-half as likely for African Americans (odds ratio=0.54, 95% CI=0.36–0.80) and Asian American/Pacific Islanders (odds ratio=0.48, 95% CI=0.27–0.85) than for non-Hispanic whites. Race/ethnicity did not contribute to the likelihood of use of 24-hour-care or informal services after the effects of other potentially confounding variables were controlled.

Table 3 presents the results of the regression analyses. The racial/ethnic group variables added significantly to the prediction of likelihood of use of outpatient services and any mental health service, after the effects of all other potentially confounding variables were taken into account. Specifically, the last column indicates that African Americans and Asian American/Pacific Islanders were approximately one-half as likely to receive any mental health service, compared to non-Hispanic whites (African Americans: odds ratio=0.56, 95% confidence interval [CI]=0.37–0.84; Asian American/Pacific Islanders: odds ratio=0.44, 95% CI=0.25–0.77). Outpatient service use was also approximately one-half as likely for African Americans (odds ratio=0.54, 95% CI=0.36–0.80) and Asian American/Pacific Islanders (odds ratio=0.48, 95% CI=0.27–0.85) than for non-Hispanic whites. Race/ethnicity did not contribute to the likelihood of use of 24-hour-care or informal services after the effects of other potentially confounding variables were controlled.

Table 1. Past-Year Mental Health Service Utilization by Youths Age 6–18 Years in a Large, Publicly Funded System of Care by Racial/Ethnic Group (N=1,256) a

<table>
<thead>
<tr>
<th>Service</th>
<th>Non-Hispanic White (N=554)</th>
<th>Latino American (N=332)</th>
<th>African American (N=282)</th>
<th>Asian American/ Pacific Islander (N=88)</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outpatient services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>χ²</td>
</tr>
<tr>
<td>Specialty mental health services b</td>
<td>67.6</td>
<td>53.4</td>
<td>51.8</td>
<td>32.8</td>
<td>0.001</td>
</tr>
<tr>
<td>Other nonspecialty services c</td>
<td>20.2</td>
<td>15.6</td>
<td>16.5</td>
<td>19.6</td>
<td>0.001</td>
</tr>
<tr>
<td>Alcohol and drug abuse treatment services</td>
<td>5.4</td>
<td>3.7</td>
<td>1.2</td>
<td>11.3</td>
<td>0.001</td>
</tr>
<tr>
<td>Any outpatient service</td>
<td>73.5</td>
<td>60.3</td>
<td>57.1</td>
<td>51.5</td>
<td>0.001</td>
</tr>
<tr>
<td>24-hour-care services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inpatient psychiatric hospital or inpatient psychiatric unit services</td>
<td>6.5</td>
<td>4.6</td>
<td>2.2</td>
<td>3.6</td>
<td>0.02</td>
</tr>
<tr>
<td>Residential treatment center or group home services</td>
<td>7.4</td>
<td>5.9</td>
<td>7.4</td>
<td>1.9</td>
<td>0.38</td>
</tr>
<tr>
<td>Alcohol or drug abuse treatment d</td>
<td>3.5</td>
<td>4.4</td>
<td>1.2</td>
<td>2.0</td>
<td>0.30</td>
</tr>
<tr>
<td>Any 24-hour-care service</td>
<td>14.4</td>
<td>14.0</td>
<td>10.2</td>
<td>6.8</td>
<td>0.23</td>
</tr>
<tr>
<td>Informal services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-help groups, peer counseling</td>
<td>17.5</td>
<td>18.5</td>
<td>7.4</td>
<td>16.7</td>
<td>0.02</td>
</tr>
<tr>
<td>Counseling from clergy</td>
<td>4.1</td>
<td>5.7</td>
<td>4.4</td>
<td>8.6</td>
<td>0.39</td>
</tr>
<tr>
<td>Alternative healer</td>
<td>1.0</td>
<td>1.9</td>
<td>0.0</td>
<td>1.0</td>
<td>0.19</td>
</tr>
<tr>
<td>Any informal service</td>
<td>20.3</td>
<td>23.5</td>
<td>11.4</td>
<td>21.2</td>
<td>0.02</td>
</tr>
<tr>
<td>Any mental health service e</td>
<td>79.0</td>
<td>70.1</td>
<td>63.7</td>
<td>59.3</td>
<td>0.001</td>
</tr>
</tbody>
</table>

a Youths were identified from the active services rolls of one or more of five San Diego County public sectors of care (alcohol and drug abuse services, child welfare services, juvenile justice services, mental health services, and public school special education services for youths with serious emotional disturbance) during the first half of 1997.
b Includes visits to professional psychologists, counselors, community mental health clinics, and partial hospitalization or day treatment programs.
c Includes visits to pediatricians and physicians and in-home therapy or emergency room visits for emotional or behavioral reasons.

d Includes treatment in an inpatient setting or residential treatment center for substance abuse problems.

e Includes all outpatient, 24-hour-care, and informal service types listed in the table.
special education services for youths with serious emotional disturbance) had very high rates of mental health service use. Nearly three-quarters of the participants used some mental health service within the past year, compared to utilization estimates for community samples ranging from 6% to 21% (2, 24). High utilization rates would be expected in this high-risk sample of youths identified from public service sectors. However, there were significantly different rates of service use across racial/ethnic groups; 79% of non-Hispanic white youths received a mental health service, compared to 59% of Asian American/Pacific Islanders, 64% of African Americans, and 70% of Latino Americans. After the effects of potentially confounding variables were controlled, the youths’ race/ethnicity was still a significant predictor of any service use, and African American and Asian American/Pacific Islanders were approximately one-half as likely as non-Hispanic white youths to receive any service.

Significant racial/ethnic differences were found for only certain types of services. Outpatient services were used most frequently overall, and there was a significant racial/ethnic group difference for this category of service use. After the effects of other predictors were controlled, African Americans and Asian American/Pacific Islanders were less likely to use outpatient services, compared to non-Hispanic whites. Latino American youths also had lower rates of outpatient service use, compared to non-Hispanic white youths, but the statistical significance of this effect was reduced by the inclusion of other predictor variables. Within the broad category of outpatient services, there were significant bivariate racial/ethnic group differences in utilization rates for specialty mental health and alcohol and drug abuse services but not for other nonspecialty care, which included pediatric visits. Previous research has also found no racial/ethnic difference in receipt of mental health services in primary care pediatric offices (29).

Thirteen percent of the participants used some 24-hour-care service in the past year, including psychiatric hospitalization, alcohol and drug abuse treatment, or residential treatment. There was no overall effect of race/ethnicity on likelihood of use of this category of services, either before or after the effects of other predictors were controlled. There was a significant bivariate group difference in use of psychiatric hospitalization, with non-Hispanic whites reporting the highest rate of use, but this ef-

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Weighted Percent</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-Hispanic White (N=554)</td>
<td>Latino American (N=332)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–11</td>
<td>37.1</td>
<td>18.7</td>
</tr>
<tr>
<td>12–15</td>
<td>28.3</td>
<td>30.3</td>
</tr>
<tr>
<td>16–18</td>
<td>34.6</td>
<td>51.0</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>32.6</td>
<td>33.5</td>
</tr>
<tr>
<td>Male</td>
<td>67.4</td>
<td>66.5</td>
</tr>
<tr>
<td>Family income by quartile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First (≤$13,000/year)</td>
<td>24.9</td>
<td>44.0</td>
</tr>
<tr>
<td>Second (≤$25,000/year)</td>
<td>23.0</td>
<td>25.5</td>
</tr>
<tr>
<td>Third (≤$45,000/year)</td>
<td>24.3</td>
<td>18.4</td>
</tr>
<tr>
<td>Fourth (&gt; $45,000/year)</td>
<td>27.8</td>
<td>12.1</td>
</tr>
<tr>
<td>Any college education</td>
<td>41.6</td>
<td>10.1</td>
</tr>
<tr>
<td>Service sector affiliation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol and drug abuse services, mental health services, or public school–based services</td>
<td>68.2</td>
<td>59.5</td>
</tr>
<tr>
<td>Child welfare services or juvenile justice services</td>
<td>31.8</td>
<td>40.5</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any DSM-IV diagnosis</td>
<td>62.7</td>
<td>48.3</td>
</tr>
<tr>
<td>Any Composite International Diagnostic Interview substance abuse module diagnosis</td>
<td>15.5</td>
<td>20.6</td>
</tr>
<tr>
<td>Functional impairment by either Columbia Impairment Scale or Children’s Global Assessment Scale cutoffs</td>
<td>67.5</td>
<td>63.2</td>
</tr>
<tr>
<td>Caregiver strain score above median</td>
<td>49.1</td>
<td>40.6</td>
</tr>
<tr>
<td>Police contact, lifetime</td>
<td>41.7</td>
<td>52.5</td>
</tr>
<tr>
<td>Caregiver depression score above clinical cutoff point (Center for Epidemiologic Studies Depression Scale)</td>
<td>28.8</td>
<td>32.2</td>
</tr>
</tbody>
</table>

*Youths were identified from the active services rolls of one or more of five San Diego County public sectors of care (alcohol and drug abuse services, child welfare services, juvenile justice services, mental health services, and public school special education services for youths with serious emotional disturbance) during the first half of 1997.*
Am J Psychiatry 162:7, July 2005

TABLE 3. Association of Mental Health Service Use With Model Variables Among Youths Age 6–18 Years in a Large, Publicly Funded System of Care (N=1,256)*

<table>
<thead>
<tr>
<th>Model Variable</th>
<th>Outpatient Services</th>
<th>24-Hour-Care Services</th>
<th>Informal Services</th>
<th>Any Mental Health Service</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio</td>
<td>CI</td>
<td>Odds Ratio</td>
<td>CI</td>
</tr>
<tr>
<td>Potentially confounding variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.98</td>
<td>0.93–1.0</td>
<td>1.1</td>
<td>0.98–1.2</td>
</tr>
<tr>
<td>Female sexb</td>
<td>1.7**</td>
<td>1.2–2.4</td>
<td>1.6*</td>
<td>1.1–2.5</td>
</tr>
<tr>
<td>Family/social factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caregiver strain</td>
<td>1.8**</td>
<td>1.2–2.7</td>
<td>2.8***</td>
<td>1.6–5.0</td>
</tr>
<tr>
<td>Caregiver depression</td>
<td>1.1</td>
<td>0.74–1.5</td>
<td>0.86</td>
<td>0.54–1.4</td>
</tr>
<tr>
<td>Police contact</td>
<td>0.54**</td>
<td>0.37–0.79</td>
<td>1.3</td>
<td>0.81–2.1</td>
</tr>
<tr>
<td>Recruitment from alcohol and drug abuse services, mental health services, or public school-based services</td>
<td>1.4</td>
<td>0.99–1.9</td>
<td>7.4***</td>
<td>3.3–16.6</td>
</tr>
<tr>
<td>Family resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incomec</td>
<td>0.69</td>
<td>0.42–1.1</td>
<td>0.70</td>
<td>0.38–1.3</td>
</tr>
<tr>
<td>Second quartile (≤$25,000/year)</td>
<td>0.51**</td>
<td>0.32–0.82</td>
<td>0.57</td>
<td>0.29–1.1</td>
</tr>
<tr>
<td>Third quartile (≥$45,000/year)</td>
<td>0.50**</td>
<td>0.31–0.83</td>
<td>0.52*</td>
<td>0.27–0.97</td>
</tr>
<tr>
<td>Any college education Insurancef</td>
<td>1.2</td>
<td>0.81–1.7</td>
<td>0.77</td>
<td>0.48–1.2</td>
</tr>
<tr>
<td>None</td>
<td>0.89</td>
<td>0.43–1.8</td>
<td>0.59</td>
<td>0.23–1.5</td>
</tr>
<tr>
<td>Government (e.g., MediCal)</td>
<td>1.1</td>
<td>0.72–1.6</td>
<td>0.58*</td>
<td>0.34–0.99</td>
</tr>
<tr>
<td>Diagnosis and impairment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any Diagnostic Interview</td>
<td>1.7**</td>
<td>1.1–2.5</td>
<td>0.88</td>
<td>0.48–1.6</td>
</tr>
<tr>
<td>Schedule for Children diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any Composite International</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic Interview substance abuse module diagnosis</td>
<td>0.90</td>
<td>0.54–1.5</td>
<td>1.3</td>
<td>0.73–2.4</td>
</tr>
<tr>
<td>Functional impairment</td>
<td>1.4</td>
<td>0.95–2.0</td>
<td>1.2</td>
<td>0.61–2.5</td>
</tr>
<tr>
<td>Logistic regression analysis for potentially confounding variables</td>
<td>6.8</td>
<td>&lt;0.0001</td>
<td>5.3</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Variable</th>
<th>Odds Ratio</th>
<th>CI</th>
<th>Odds Ratio</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race/ethnic group†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino American</td>
<td>0.71</td>
<td>0.46–1.1</td>
<td>1.0</td>
<td>0.57–1.8</td>
</tr>
<tr>
<td>African American</td>
<td>0.54**</td>
<td>0.36–0.80</td>
<td>1.1</td>
<td>0.61–1.9</td>
</tr>
<tr>
<td>Asian American/Pacific Islander</td>
<td>0.48*</td>
<td>0.27–0.85</td>
<td>0.40*</td>
<td>0.17–0.96</td>
</tr>
<tr>
<td>Logistic regression analysis for race/ethnic group variable</td>
<td>4.2</td>
<td>0.0006</td>
<td>1.6</td>
<td>0.19</td>
</tr>
<tr>
<td>Logistic regression analysis for all variables</td>
<td>6.3</td>
<td>&lt;0.0001</td>
<td>5.0</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

* Youths were identified from the active services rolls of one or more of five San Diego County public sectors of care (alcohol and drug abuse services, child welfare services, juvenile justice services, mental health services, and public school special education services for youths with serious emotional disturbance) during the first half of 1997.

b Reference group: male participants.

c Reference group: families with income in the fourth quartile (>$45,000/year).

d Reference group: participants with private insurance.

e Reference group: non-Hispanic white participants.

*<0.05. **p<0.01. ***p<0.001.

f Effects were not statistically significant when the effects of other predictor variables were controlled.

Approximately 20% of the participants used informal services, including self-help or peer counseling groups, counseling from clergy, and alternative healers. There was a significant racial/ethnic group difference in use of this category of services, with Latino American youths exhibiting the highest rates of use and African American youths the lowest. However, after the effects of other predictors were controlled, there was no racial/ethnic group difference. Given that the racial/ethnic minority participants did not use these services significantly more frequently than did non-Hispanic whites, there was no evidence that higher rates of informal service use among minority groups compensated directly for lower rates of use of professional services. Previous research with adults in the community suggested that Latino Americans with mental health problems often utilize informal or nonspecialty services (30).

The findings regarding the other (i.e., nonrace/ethnicity) potentially confounding predictors of any mental health service use were generally consistent with previous findings for youths in the community. Positive predictors in-
cluded diagnosis and impairment, caregiver strain, female gender, and involvement with the alcohol or drug abuse treatment, mental health, or school-based service sectors; negative predictors included moderate family income and police contact (5, 17, 31). The only factor associated with all types of service use was parental report of caregiver strain. Previous research has supported the importance of this factor in youth mental health service use (7, 12, 13, 31), and significant racial/ethnic differences in caregiver strain have been identified, with non-Hispanic whites generally reporting higher levels of strain, compared to other groups (14). These results confirm the strength of the association between caregiver strain and service use but indicate that racial/ethnic differences in reported strain do not entirely explain differences in service utilization.

Contact with the police and/or juvenile justice system was associated with a significantly lower likelihood of professional outpatient service use but a higher likelihood of informal service use. The high prevalence of substance abuse problems among these youths is likely to be associated with referral to Alcoholics Anonymous–related services (8, 18). Referral rates from the justice system to formal mental health services are much lower than would be expected, given the high level of need for mental health services in this population (15).

**Limitations**

This study addressed many of the limitations of previous work in this area by 1) including sufficient representation of youths from four major racial/ethnic groups; 2) using well-established instruments to assess service use, psychiatric diagnoses, and functional impairment; 3) including assessment of most of the other factors known to predict service use, such as caregiver strain, police contact, and family income; and 4) assessing use of a variety of types of mental health services. There are, however, some limitations. The service use data relied on parents’ and youths’ self-reports, which may be subject to cognitive and culturally influenced biases. Specifically, there may be cultural differences in comfort in disclosing mental health service use and/or labeling of service use. In addition, commonly used school-based services were not included in these analyses. Also, the participation rate was less than optimal (67%), but retrospective analyses did not identify major differences in the basic sociodemographic characteristics of participants, compared to nonparticipants. Although the sample was diverse, it excluded non-English-speaking families and thus may not be representative of all families in each of these four major racial/ethnic groups. Non-English speakers are likely to be less acculturated to the dominant society and less likely to utilize mental health services (32). Thus, exclusion of these families may have actually underestimated the service use disparities for some racial/ethnic minority groups. Finally, the sample was a “hybrid,” high-risk sample of youths who had some contact with the public system of care, and thus the findings do not generalize to general community samples.

The present study was limited to a cross-sectional analysis of self-reported mental health service use and could not elucidate the causes of the racial/ethnic disparities that were identified. More research is needed to examine the dynamic and complex processes that lead to youth mental health service utilization and the specific reasons for these racial/ethnic disparities. The network/episode model of access to care, adapted for children and adolescents, suggests that utilization of services is affected by multiple interacting factors, ranging from individual (child and family) help-seeking preferences to broader system-level factors, including access, availability, referral practices, and funding policies (33). In terms of individual characteristics, culturally influenced cognitive explanatory models for constructs such as problem recognition, etiological attribution, and relevance of mental health services have been hypothesized to help explain racial/ethnic differences in service use in adult populations (34, 35). In support of this theory for youth populations, racial/ethnic differences in parental beliefs about children’s behavioral problems have been identified (36, 37), and these differences may partly explain service use patterns.

More research is needed on referral practices and potential biases among “gateway providers” (e.g., case workers, probation officers, school counselors, teachers) in the identification of the mental health needs of youths and referral to services (38). Interdisciplinary approaches are needed to support investigation of how organizational, policy, and funding factors interact with community and family help-seeking preferences to result in differential access and utilization. Newer models of mental health help seeking, such as the network/episode model (33), which include greater attention to the social context, may be very useful in elucidating the contextual factors that influence individual family decisions about service utilization.

**Conclusions**

Racial/ethnic disparities in use of professional mental health services are robust. The results of this study reinforce the public’s and policy makers’ concerns about racial/ethnic disparities in mental health service utilization for vulnerable youths and families.

Received April 1, 2003; revisions received Nov. 25, 2003, April 7 and May 18, 2004; accepted June 14, 2004. From Child and Adolescent Services Research Center; the Department of Psychiatry, University of California, San Diego, La Jolla, Calif.; the Department of Psychology, University of California, Los Angeles, Los Angeles; the School of Social Work and the Department of Psychology, San Diego State University, San Diego; the Department of Psychology, University of San Diego, San Diego; and the Department of Psychiatry, University of New Mexico, Albuquerque. Address correspondence and reprint requests to Dr. Garland, Child and Adolescent Services Research Center, Children’s Hospital and Health Center, 3020 Children’s Way (MC 5033), San Diego, CA 92123; agarland@casrc.org (e-mail).

The Patterns of Youth Mental Health Care in Public Service Systems study is supported by NIMH grant MH-55282. Preparation of this ar...
References

15. Breda C: The mental health orientation of juvenile courts. J Behav Health Serv Res 2001; 28:89–95

http://ajp.psychiatryonline.org

Am J Psychiatry 162:7, July 2005

GARLAND, LAU, YEH, ET AL.
A 3-Year Panel Study of Mental Disorders Among Adolescents in Taiwan

Susan S.F. Gau, M.D., Ph.D.
M.Y. Chong, M.D., Ph.D.
Tony H.H. Chen, M.D., Ph.D.
Andrew T.A. Cheng, M.D., Ph.D., D.Sc., F.R.C.Psych.

Objective: This study investigated the prevalence and changing trends of mental disorders and the effects of gender and urbanization among adolescents in Taiwan.

Method: A random sample of seventh-grade students (N=1,070) was recruited from one urban and one rural junior high school in which 1,051 (98.2%) and 1,035 (96.7%) were reassessed in the second and third years, respectively. A two-stage case identification was conducted by mental health professionals with the Schedule for Affective Disorders and Schizophrenia for School-Age Children—Epidemiologic Version, supplemented by information from the Child Behavior Checklist.

Results: The weighted 3-month prevalence rates across the 3 consecutive years for overall psychiatric disorders were 20.3%, 22.7%, and 14.8%, respectively. The most prevalent psychiatric condition was attention deficit hyperactivity disorder (ADHD) in the first 2 years and substance use disorders in the third. During the 3 years, the rates for ADHD, specific phobia, and social phobia decreased, and the rates for major depression and substance use disorders, conversely, increased. Although conduct disorder, ADHD, and substance use disorders were more prevalent among boys, the rates for major depression, social phobia, specific phobia, and adjustment disorder were higher among girls. Rural adolescents had higher rates of conduct disorder, oppositional defiant disorder, and substance use disorders than their urban counterparts.

Conclusions: Our findings are similar to those of previous studies among adolescents in prevalence rates, changing trends of most mental disorders, and gender effects. The differential changing trends in various diagnostic groups may imply the importance of specific measures for prevention during adolescence.
substance use disorders (9). The two-stage case finding strategy was applied with a brief screening tool for any substance use disorders and the Chinese version of the Schedule for Affective Disorders and Schizophrenia for School-Age Children—Epidemiologic Version (K-SADS-E) (9, 24, 25). The lifetime weighted prevalence rate of any DSM-III-R substance use disorder was 11.0% and was significantly higher in boys, in a rural community, and in classes with poor academic performance (9). A 3-year panel study was then carried out to investigate the prevalence and trend of psychiatric disorders in 1995–1997. This article reports the research methods, prevalence rates of specific mental disorders, and their changing trends in relation to gender and urban or rural settings across the 3 years.

Method

This was a school-based panel study among a sample of seventh-grade students established in the first year of the survey.

Study Subjects

Eighteen of 34 classes and eight of 10 classes in grade 7 were randomly selected from one urban and one rural junior high school in South Taiwan, respectively. The two schools were selected based on the positive response and cooperativeness of their principals. The former represented a typical urban school with a high pass rate in the joint entrance examination for senior high school; the latter represented a typical rural school with a low pass rate in such an examination over the years. All the students in the selected classes were included in the study sample. The study sample consisted of 1,070 students (532 boys and 538 girls, ages 13–15), in which 725 (67.8%) were from urban areas (357 boys and 368 girls) and 345 (32.2%) were from rural areas (175 boys and 170 girls). The follow-up rates were 98.2% (N=1,051) and 96.7% (N=1,035) in the second year (eighth grade) and the third year (ninth grade), respectively. No significant differences in gender and urban-rural distribution were found between the respondents and the nonrespondents.

Measures

Chinese K-SADS-E. The K-SADS-E is a semistructured clinical interview for the systematic assessment of both past and current episodes of mental disorders in children and adolescents (24). Development of the Chinese version of the K-SADS-E was carried out by the Child Psychiatry Research Group in Taiwan (25), which included a two-stage translation and modification of several items with psycholinguistic equivalents relevant to Taiwanese culture. Further modification to meet the DSM-IV diagnostic criteria and an additional section developed for betel use disorder were performed by the research team (9). In this panel study, all the screening items in individual sections of the Chinese K-SADS-E were grouped together to form a separate screening version for use in the first stage of case finding.

The interrater reliability of the Chinese K-SADS-E was examined before administration of the cross-sectional survey among nine consultant child psychiatrists in the research team, and the results showed generalized kappa coefficients ranging from 0.73 to 0.96 for all mental disorders included in the Chinese K-SADS-E. Validity of the screening version of the Chinese K-SADS-E was assessed before the administration of this panel study among 124 randomly selected eighth-grade students. The screening interview was first carried out by 14 psychiatric clinical staff members (nurses, psychologists, social workers, and psychiatric residents), who were instructed to record the presence/absence of all the screening items entirely, according to the response from the study subjects without any personal judgment. Based on the Chinese K-SADS-E interview, the nine consultant child psychiatrists then made psychiatric diagnoses. The subjects with a positive response to any of the Chinese K-SADS-E screening items were treated as screened positive. In this pretest, the overall sensitivity and specificity of the screening interview against any Chinese K-SADS-E diagnostic category were calculated to be 78% and 98%, respectively.

Teacher report form of the Child Behavior Checklist. The Chinese version of the teacher report form of the Child Behavior Checklist has been proved to be a reliable and valid instrument when used in Taiwanese adolescent populations (26).

The Fieldwork

The Review Board of the Department of Health, Taiwan, approved this study as ethical for studying adolescents. Child assent and oral informed consent were obtained from the study subjects and their parents, respectively, after detailed explanation of the purpose and interview procedures of this study, and confidentiality about interview records was ensured. The fieldwork was then conducted at school following a timetable arranged by tutors of the study classes.

The same threshold used in the pretest for the first-stage screening was employed in the main survey. In the second stage, all of those who were screened positive and every 1 in 10 who was randomly selected from those who were screened negative were immediately given the second-stage Chinese K-SADS-E interview, conducted by child psychiatrists who were blind to the screening results. There was no time lag between the first- and second-stage interviews, and none of the respondents who received the screening interview refused to take the diagnostic interview. The same two-stage case-finding procedure was conducted in the next 2 consecutive years. Teacher report forms were collected soon after the fieldwork.

Psychiatric Diagnosis

A psychiatric diagnosis of the study subjects was first made by child psychiatrists who conducted the Chinese K-SADS-E interview according to the DSM-IV. These diagnoses were then independently reassessed by two senior psychiatrists (S.S.F.G. and A.T.A.C.) by a systematic review of all the interview records. In our reassessment, the principle of rate-down was employed, and any information that was dubious or uncertain was discarded. Psychiatric diagnoses generated from this reassessment were jointly discussed, and our consensus diagnosis was taken as final. To minimize the likely underreporting of externalized disorders by the study subjects, information regarding behavioral syndromes gathered from the teacher report forms was incorporated into our diagnostic consideration for the best estimation of ADHD, oppositional defiant disorder, and conduct disorder.

Statistical Analysis

Statistical analyses were performed with SAS, version 8.2 (SAS Institute, Cary, N.C.). The preselected alpha value was 0.05. The 3-month weighted prevalence rates and their variance for individual psychiatric disorders were calculated with the formula

$$p = \frac{\lambda_1 + \lambda_2 (1 - \pi)}{N}$$

and

$$V \left( \frac{p}{N(1 - \pi)} \right) = \frac{1}{N} \left[ \pi (1 - \pi) \left( \lambda_1 + \lambda_2 \right) \right],$$

where $p$ was the weighted prevalence, $\pi$ was the fraction of the sample screened positive, $1 - \pi$ was the proportion of the sample screened negative, $f$ was the fraction of the sample screened negative who were interviewed at the second stage, $\lambda_1$ was the proportion of cases among the sample screened positive who were interviewed, and $\lambda_2$ was the proportion of cases among the sam-
MENTAL DISORDERS AMONG ADOLESCENTS

TABLE 1. Diagnostic Distribution of DSM-IV Psychiatric Disorders in an Adolescent Sample in Taiwan

<table>
<thead>
<tr>
<th>DSM-IV Diagnoses</th>
<th>Seventh Grade (N=1,070)</th>
<th>Eighth Grade (N=1,051)</th>
<th>Ninth Grade (N=1,035)</th>
<th>z&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>95% CI&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N</td>
</tr>
<tr>
<td>Disruptive behavioral disorders</td>
<td>105</td>
<td>9.8</td>
<td>7.2 to 12.4</td>
<td>94</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>27</td>
<td>2.5</td>
<td>1.6 to 3.5</td>
<td>30</td>
</tr>
<tr>
<td>Oppositional defiant disorder</td>
<td>20</td>
<td>1.9</td>
<td>1.1 to 2.7</td>
<td>27</td>
</tr>
<tr>
<td>Attention deficit hyperactivity disorder</td>
<td>81</td>
<td>7.5</td>
<td>5.1 to 10.0</td>
<td>64</td>
</tr>
</tbody>
</table>

Depressive disorders
- Any depressive disorder: 7, 0.7, 0.2 to 1.1
- Major depression: 5, 0.5, 0.1 to 0.9
- Dysphoric disorder: 2, 0.2, -0.1 to 0.4

Anxiety disorders
- Any anxiety disorder: 99, 9.2, 6.1 to 12.3
- Generalized anxiety disorder: 7, 0.7, 0.2 to 1.1
- Social phobia: 36, 3.4, 1.3 to 3.0
- Specific phobia: 54, 5.0, 2.8 to 7.3
- Separation anxiety disorder: 3, 0.3, 0.0 to 0.6
- Panic disorder: 2, 0.2, -0.1 to 0.4
- Agoraphobia without panic: 2, 0.2, -0.1 to 0.4
- Obsessive-compulsive disorder: 2, 0.2, -0.1 to 0.4

Adjustment disorders: 14, 1.3, -0.7 to 3.3

Substance use disorders
- Any substance use disorders: 24, 2.2, 1.4 to 3.1
- Nicotine use disorders
  - Any nicotine use disorders: 24, 2.2, 1.4 to 3.1
  - Nicotine-related disorder: 18, 1.7, 0.9 to 2.5
  - Nicotine dependence: 6, 0.6, 0.1 to 1.0
- Betel use disorders
  - Any betel use disorder: 9, 0.8, 0.3 to 1.4
  - Betel abuse: 8, 0.8, 0.2 to 1.3
  - Betel dependence: 1, 0.1, -0.1 to 0.3
- Alcohol use disorders
  - Any alcohol use disorder: 5, 0.5, 0.1 to 0.9
  - Alcohol abuse: 5, 0.5, 0.1 to 0.9
  - Alcohol dependence: 0, 0.0, 0.0 to 0.0
- One substance use disorder: 13, 1.2, 0.6 to 1.9
- Two or more substance use disorders: 11, 1.0, 0.4 to 1.6
- Any psychiatric disorder: 217, 20.3, 15.9 to 24.6

<sup>a</sup>Cochran-Armitage trend test.
<sup>b</sup>Weighted prevalence.
*<sup>p</sup><0.05. **<sup>p</sup><0.01. ***<sup>p</sup><0.001. †<sup>p</sup><0.0001.

Results

Prevalence of Psychiatric Disorders
Table 1 shows the DSM-IV diagnostic distribution of all panel respondents from seventh to ninth grade with weighted 3-month prevalence rates and their 95% CIs. The overall rates for the seventh and eighth grades were higher than that for the ninth, with a significant decline in trend. Such decline was mainly observed for disruptive behavioral disorders and anxiety disorders, specifically among child-onset disorders (including ADHD, specific phobia, social phobia, and separation anxiety disorder).

No such trend was observed for conduct disorder and oppositional defiant disorder. Both depressive disorders (notably major depression) and substance use disorders conversely increased across the three grades. Moreover, the increasing trend was significant, mainly for comorbid substance use disorders and not for any single substance use disorder.

In grades 7 and 8, the most prevalent diagnostic group was disruptive behavioral disorders (notably ADHD), followed by anxiety disorders (notably specific phobia), and substance use disorders. In grade 9, although disruptive behavioral disorders was still ahead of others, the rate for substance use disorders increased to become the second leading disorder, and depressive disorders (mainly major depression) increased to become third. Rates for the top three disorders in grade 9 were, in fact, close to each other. In all three grades, the most common substance of abuse/dependence was nicotine, followed by betel and alcohol.

Gender Difference
Disruptive behavioral disorders (mainly conduct disorder and ADHD) and substance use disorders were more...
prevalent in boys than in girls, whereas the reverse was observed for depressive disorders across the 3 years (Table 2). Girls had higher rates of anxiety disorders and adjustment disorder than boys, particularly, higher rates of specific phobia in grade 8 and social phobia in grade 9.

The significant changing trends in this panel for most mental disorders over the 3 years were observed in boys and girls. However, the increasing trend for all substance use disorders was only significant in boys. The declining trend for social phobia was only significant in boys and that for adjustment disorder was only significant in girls.

**Urban-Rural Difference**

The overall rates of mental disorders were generally higher in rural than in urban youths (Table 3). Compared to their urban counterparts, rural adolescents had significantly higher rates for conduct disorder, oppositional defiant disorder, and substance use disorders over the 3 years. In grade 9, specific phobia was more prevalent among rural adolescents; in contrast, social phobia was more common among urban adolescents. There was no urban-rural difference in rates for depressive and other anxiety disorders in all grades.

The time trends for most disorders in rural and urban areas also followed those in the total sample, with some exceptions. The increasing trend for all substance use disorders was significant in rural areas but was less apparent in urban areas and only significant for nicotine use disorders. The declining trend for social phobia was only significant among rural adolescents and that for specific phobia was greater in urban than in rural areas, notably in grade 9.

**TABLE 2. Gender Differences in Psychiatric Morbidity Across Three Grades in an Adolescent Sample in Taiwan**

<table>
<thead>
<tr>
<th>DSM-IV Diagnoses</th>
<th>Seventh-Grade Boys (N=532) Versus Girls (N=538)</th>
<th>Eighth-Grade Boys (N=526) Versus Girls (N=525)</th>
<th>Ninth-Grade Boys (N=517) Versus Girls (N=518)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio 95% CI</td>
<td>Odds Ratio 95% CI</td>
<td>Odds Ratio 95% CI</td>
</tr>
<tr>
<td>Disruptive behavior disorders</td>
<td>3.4 2.1 to 5.4†</td>
<td>4.3 2.6 to 7.1†</td>
<td>4.2 2.2 to 8.1†</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>4.6 1.7 to 15.6**</td>
<td>5.2 1.9 to 17.5**</td>
<td>6.8 2.3 to 27.0†</td>
</tr>
<tr>
<td>Oppositional defiant disorder</td>
<td>1.9 0.8 to 4.8</td>
<td>2.7 1.2 to 6.1°</td>
<td>1.6 0.5 to 6.3°</td>
</tr>
<tr>
<td>Attention deficit hyperactivity disorder</td>
<td>4.5 2.5 to 7.8†</td>
<td>5.2 2.7 to 10.1**</td>
<td>4.9 2.0 to 11.9†</td>
</tr>
<tr>
<td>Depressive disorders</td>
<td>0.1 0.0 to 1.2**</td>
<td>0.2 0.0 to 0.5**</td>
<td>0.5 0.3 to 1.0†</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any anxiety disorder</td>
<td>1.0 0.6 to 1.4</td>
<td>0.3 0.2 to 0.4†</td>
<td>0.3 0.1 to 0.6†</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>0.4 0.0 to 2.5</td>
<td>0.5 0.0 to 9.6†</td>
<td>—</td>
</tr>
<tr>
<td>Social phobia</td>
<td>1.1 0.6 to 2.2</td>
<td>0.5 0.2 to 1.2</td>
<td>0.3 0.1 to 0.7†</td>
</tr>
<tr>
<td>Specific phobia</td>
<td>0.8 0.5 to 1.4</td>
<td>0.2 0.1 to 0.4†</td>
<td>1.3 0.2 to 9.2†</td>
</tr>
<tr>
<td>Adjustment disorder</td>
<td>0.2 0.0 to 0.8**</td>
<td>0.2 0.0 to 0.7**</td>
<td>—</td>
</tr>
<tr>
<td>Substance use disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any substance use disorder</td>
<td>4.0 1.4 to 13.6**</td>
<td>5.6 2.5 to 12.7†</td>
<td>5.5 2.7 to 11.4†</td>
</tr>
<tr>
<td>Nicotine use disorders</td>
<td>4.0 1.4 to 13.6**</td>
<td>6.6 2.7 to 15.6†</td>
<td>5.5 2.7 to 11.4†</td>
</tr>
<tr>
<td>Betel use disorders</td>
<td>19.5 1.1 to 336.7**</td>
<td>40.4 2.4 to 670.6†</td>
<td>60.4 3.7 to 991.6†</td>
</tr>
<tr>
<td>Alcohol use disorders</td>
<td>1.5 0.2 to 18.3†</td>
<td>5.6 1.2 to 52.0**</td>
<td>5.8 1.7 to 31.2**</td>
</tr>
<tr>
<td>Any psychiatric disorder</td>
<td>1.7 1.2 to 2.3**</td>
<td>1.1 0.8 to 1.4</td>
<td>1.5 1.1 to 2.1†</td>
</tr>
</tbody>
</table>

* p values are for comparisons between boys and girls by chi-square or Fisher’s exact test.
† Fisher’s exact test.
‡ Cochran-Armitage trend test.
§ Fisher’s exact test, adding 0.5 to each cell to compute odds ratios and 95% CIs.
* p<0.05. ** p<0.01. *** p<0.001. †† p<0.0001.

**Discussion**

**Methodology Considerations**

As one of the few studies (1, 8, 28) that have investigated the prevalence and the time trend for adolescent psychiatric morbidities, the present study provided a unique opportunity to compare the sex and the urban-rural difference of such morbidities between developing and developed countries. Our study has employed the two-stage design, and all the clinical interviews were conducted by child mental health professionals with the standardized Chinese K-SADS-E interview records and a joint discussion by two senior research psychiatrists. The response rates at phase I and the two follow-ups were very high.

Despite all of these strengths, there are some limitations of this study that require careful consideration in the interpretation of the findings. First, because of a purposeful sampling of study schools, its external validity for other Taiwanese adolescent populations needs to be examined. Second, the psychiatric diagnoses were mainly based on clinical interviews of study subjects without interviewing their parents. Previous studies have shown low agreement among child, parent, and teacher informants in reporting children’s emotional and behavioral problems (29) and the need to incorporate teachers’ reports into the identification of externalizing disorders (29). Therefore, we included the teacher report form to make the best estimates of psychiatric diagnoses of ADHD, conduct disorder, and oppositional defiant disorder. Finally, in this two-stage case-finding study, despite the fact that all the study subjects


1347
were screened at the first stage for 3 consecutive years, the second-stage psychiatric interview was not performed among all of them. The lack of complete information in psychiatric diagnoses for all study subjects has prevented us from conducting longitudinal analyses using a multilevel model to examine the trajectories of psychiatric diagnoses at an individual level.

**Prevalence and Trend**

The magnitude of total psychiatric morbidities estimated in this study was in the middle of previous studies (5, 6) and was similar to those (20.9% and 20.3%) of two large-scale epidemiology studies of youths in the United States (4, 8). However, the clinical interview and diagnostic criteria (DSM-III-R) used in those two U.S. studies were different from those in this study. Shaffer et al. (4) employed the National Institute of Mental Health's Diagnostic Interview Schedule for Children, Version 2.3, and Costello et al. (7) used the Child and Adolescent Psychiatric Assessment.

Although there was only a minor change in the diagnostic criteria for ADHD from DSM-III-R to DSM-IV, the average estimated prevalence of ADHD has been reported to increase from 3%–5% with DSM-III-R to 9%–10% with DSM-IV with the three newly created subtypes (11, 30). The overall prevalence of ADHD in this study was close to the figures in recent studies in Australia (31) with the Diagnostic Interview Schedule for Children and in Brazil (32) with the 18 DSM-IV ADHD symptom items and clinical diagnosis. Similar to findings in previous studies (15), the rates of ADHD declined over adolescence. Our rates of conduct disorder and oppositional defiant disorder were in the lower part of the reported rates across cultures and countries (8, 33), with a nonsignificant increase for conduct disorder over the 3 years in boys.

A unique finding of this study is that unlike Western societies (19), betel instead of alcohol was the second (after nicotine) most prevalent abused substance among our study subjects, which might be attributable to an increased availability of betel because of commercial interest in recent years and the popularity of the betel chewing habit, particularly in rural areas (9). The fact that all the study subjects with substance use disorders had nicotine use disorder and the rates of nicotine dependence had markedly increased over time strongly indicates the seriousness of nicotine damage on health among Taiwanese adolescents today.

The 3-month prevalence rate of depressive disorders in this study was somewhat lower than those in previous studies, which largely reported 6- or 12-month prevalence rates in similar age populations (4, 8). However, our finding of a significant increase in time trend of such morbidity in both sexes was consistent with previous work (20).

Consistent with previous studies, specific phobia was the most common anxiety disorder, followed by social phobia, then generalized anxiety disorder (12–14). The lower trend in the rates of total anxiety disorders mainly came from specific phobia, social phobia, and separation anxiety disorder (16), and girls had a greater stability of internalized disorders than boys over time (34).

**Gender Difference**

Similar to earlier studies, we found higher rates of ADHD, conduct disorder, oppositional defiant disorder, and substance use disorders in boys, contributing to a higher total psychiatric morbidity than girls. Higher rates of anxiety

<table>
<thead>
<tr>
<th>DSM-IV Diagnosis</th>
<th>Seventh-Grade Adolescents: Rural (N=345) Versus Urban (N=725)</th>
<th>Eighth-Grade Adolescents: Rural (N=342) Versus Urban (N=709)</th>
<th>Ninth-Grade Adolescents: Rural (N=339) Versus Urban (N=696)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio 95% CI</td>
<td>Odds Ratio 95% CI</td>
<td>Odds Ratio 95% CI</td>
</tr>
<tr>
<td>Disruptive behavioral disorders</td>
<td>1.5 1.0 to 2.2</td>
<td>1.4 0.9 to 2.2</td>
<td>1.8 1.1 to 3.1†</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>3.2 1.5 to 6.9**</td>
<td>2.8 1.3 to 5.8**</td>
<td>4.3 2.0 to 9.3†</td>
</tr>
<tr>
<td>Oppositional defiant disorder</td>
<td>3.2 1.3 to 8.0**</td>
<td>3.3 1.5 to 7.2**</td>
<td>1.8 0.6 to 5.3</td>
</tr>
<tr>
<td>Attention deficit hyperactivity disorder</td>
<td>1.0 0.6 to 1.7</td>
<td>0.9 0.5 to 1.5</td>
<td>1.1 0.6 to 2.3</td>
</tr>
<tr>
<td>Depressive disorder</td>
<td>0.8 0.1 to 5.2†</td>
<td>0.4 0.1 to 1.1†</td>
<td>1.1 0.6 to 1.9</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any anxiety disorder</td>
<td>1.0 0.6 to 1.5</td>
<td>0.9 0.6 to 1.5</td>
<td>0.6 0.2 to 1.3‡</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>——</td>
<td>——</td>
<td>0.7 0.0 to 8.6‡</td>
</tr>
<tr>
<td>Social phobia</td>
<td>1.2 0.6 to 2.4</td>
<td>1.2 0.5 to 3.1</td>
<td>0.2 0.0 to 0.9**‡</td>
</tr>
<tr>
<td>Specific phobia</td>
<td>1.0 0.6 to 1.8</td>
<td>1.0 0.6 to 1.8</td>
<td>5.2 0.8 to 54.7†‡</td>
</tr>
<tr>
<td>Substance use disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any substance use disorder</td>
<td>15.6 4.6 to 82.0†</td>
<td>12.2 5.4 to 27.6†</td>
<td>7.4 3.9 to 14.1†‡</td>
</tr>
<tr>
<td>Nicotine use disorders</td>
<td>15.6 4.6 to 82.0†</td>
<td>11.8 5.2 to 26.8†</td>
<td>7.4 3.9 to 14.1†‡</td>
</tr>
<tr>
<td>Betel use disorders</td>
<td>41.0 2.4 to 705.9†‡</td>
<td>39.3 6.1 to 1641.0†‡</td>
<td>28.8 7.1 to 251.4†‡</td>
</tr>
<tr>
<td>Alcohol use disorders</td>
<td>8.5 0.8 to 418.8**</td>
<td>7.1 1.8 to 40.3**</td>
<td>12.2 3.5 to 65.3†‡</td>
</tr>
<tr>
<td>Any psychiatric disorder</td>
<td>1.3 0.9 to 1.7</td>
<td>1.4 1.1 to 1.9**</td>
<td>1.7 1.2 to 2.4**‡</td>
</tr>
</tbody>
</table>

a p values are for comparison between rural and urban by chi-square or Fisher's exact test.
b Cochrans-Armitage trend test.
c Fisher's exact test.
d Fisher's exact test, adding 0.5 to each cell to compute odds ratio and 95% CI.
*p<0.05. **p<0.01. ***p<0.001. †p<0.0001.
disorders in girls, notably specific phobia and social phobia (35), were also observed. The emerging trend of female excess in depressive disorders at adolescence reported in previous Western surveys (13, 20, 34, 36) was also evident in this study.

**Urban-Rural Difference**

Our findings of higher rates of conduct disorder/oppositional defiant disorder and substance use disorders in rural areas, although contrary to that in earlier Western studies (18, 33), were in accordance with those from recent studies in Western (37) and Asian societies (9). The finding of a greater increase in time trend for substance use disorders in rural areas in this study awaits further examination in other societies.

Urban-rural differences in psychiatric morbidity are likely to be associated with multiple social environmental factors. Urban neighborhood as a risk factor for psychiatric disorders has been explained by its close association with lower socioeconomic status in most cities in developed countries (38). This notion might have at least in part explained the higher rates of conduct disorder/oppositional defiant disorder, substance use disorders, and total psychiatric morbidities among adolescents in rural Taiwan, where the socioeconomic status has been generally lower than in their urban counterparts (9, 39). Another important factor for such difference in morbidity in Taiwan may have come from a positive selective migration of the mentally fit from rural to urban cities (39). For example, rural children from the families of higher socioeconomic status move to urban cities for better educational opportunities, leaving behind those who are more socioeconomically disadvantaged and less academically competent, with a higher vulnerability to both psychological disorders and substance abuse (9). In consequence, performance on the Joint Entrance Examination for Senior High School is generally better among urban junior high schools than among their rural counterparts (9).

Findings regarding the relationships between low socioeconomic status and anxiety and mood disorders have been contradictory (8, 35). The present study did not find any significant association between urban or rural residency, a proxy for socioeconomic status, and the rates of ADHD, depressive disorders, and most anxiety disorders. Our finding of different time trends for specific phobia and social phobia across urban or rural residency may deserve further inquiry. It is likely that environmental exposure may last longer for specific phobia in rural areas globally and may be higher regarding the pressure from social contact and school performance in urban areas in developing countries.

**Implications for Prevention**

It is imperious to provide a protective environment to prevent childhood-onset emotional and behavioral disorders and substance-related disorders among vulnerable adolescents. Our findings have implied that the amelioration of detrimental risk factors in social environment (be it more prevalent in rural or urban settings in different societies) for mental disorders in adolescents may serve as the target for primary prevention. Betel abuse, a specific disorder in Taiwan and certain Asian countries, ought to be prevented among both adolescent and adult populations, especially in rural areas. Further investigation of risk factors, patterns of comorbidity, and the trajectories of psychopathology during adolescence is crucial for the identification of the target for primary prevention among different vulnerable groups.

**References**


**Am J Psychiatry 162:7, July 2005**

http://ajp.psychiatryonline.org
MENTAL DISORDERS AMONG ADOLESCENTS


A Randomized, Double-Blind, Placebo-Controlled Trial of Quetiapine in the Treatment of Bipolar I or II Depression

Joseph R. Calabrese, M.D.  
Paul E. Keck, Jr., M.D.  
Wayne Macfadden, M.D.  
Margaret Minkwitz, Ph.D.  
Terence A. Ketter, M.D.  
Richard H. Weisler, M.D.  
Andrew J. Cutler, M.D.  
Robin McCoy, R.N.  
Ellis Wilson, M.S.  
Jamie Mullen, M.D.  
The BOLDER Study Group

Objective: There is a major unmet need for effective options in the treatment of bipolar depression.

Method: Five hundred forty-two outpatients with bipolar I (N=360) or II (N=182) disorder experiencing a major depressive episode (DSM-IV) were randomly assigned to 8 weeks of quetiapine (600 or 300 mg/day) or placebo. The primary efficacy measure was mean change from baseline to week 8 in the Montgomery-Åsberg Depression Rating Scale total score. Additional efficacy assessments included the Hamilton Depression Rating Scale, Clinical Global Impression of severity and improvement, Hamilton Anxiety Rating Scale, Pittsburgh Sleep Quality Index, and Quality of Life Enjoyment and Satisfaction Questionnaire.

Results: Quetiapine at either dose demonstrated statistically significant improvement in Montgomery-Åsberg Depression Rating Scale total scores compared with placebo from week 1 onward. The proportions of patients meeting response criteria (≥50% Montgomery-Åsberg Depression Rating Scale score improvement) at the final assessment in the groups taking 600 and 300 mg/day of quetiapine were 58.2% and 57.6%, respectively, versus 36.1% for placebo. The proportions of patients meeting remission criteria (Montgomery-Åsberg Depression Rating Scale ≤12) were 52.9% in the groups taking 600 and 300 mg/day of quetiapine versus 28.4% for placebo. Quetiapine at 600 and 300 mg/day significantly improved 9 of 10 and 8 of 10 Montgomery-Åsberg Depression Rating Scale items, respectively, compared to placebo, including the core symptoms of depression. Treatment-emergent mania rates were low and similar for the quetiapine and placebo groups (3.2% and 3.9%, respectively).

Conclusions: Quetiapine monotherapy is efficacious and well tolerated for the treatment of bipolar depression.

Depressive episodes in bipolar I and II disorder are an important source of morbidity and mortality. While symptomatic, patients with bipolar I disorder experience depressive symptoms for about threefold longer than manic symptoms, and the recovery time is considerably longer for depressive than manic episodes (1–4). Symptomatic patients with bipolar II disorder spend almost 40 times longer depressed than hypomanic patients (5). Bipolar depression is associated with high rates of disability (6) and an increased risk of suicide, which occurs in 10% to 20% of patients with bipolar disorder (7).

Although multiple agents, including several atypical antipsychotics, have demonstrated efficacy in the treatment of the manic phase of bipolar I disorder (8), the acute treatment of bipolar depression has not been as well studied (9). Lithium and lamotrigine are recommended as initial treatments for acute bipolar I depression (10, 11). However, the response of bipolar depression to lithium is often incomplete in a substantial proportion of patients (12), and the efficacy of lamotrigine in the treatment of acute bipolar I depression has only been demonstrated in one adequately powered placebo-controlled trial (13).

More recently, the atypical antipsychotic olanzapine was found to be superior to placebo in the treatment of acute bipolar I depression as monotherapy when data were pooled from two 8-week trials (14). Fixed doses of olanzapine in combination with the antidepressant fluoxetine were administered to small groups of patients in these studies and were found to be both superior to placebo and superior to olanzapine monotherapy.

Quetiapine is efficacious in the treatment of acute bipolar mania, both as monotherapy and in combination with other mood stabilizers (15, 16). Preliminary evidence for the efficacy of quetiapine in the treatment of depressive symptoms in a variety of psychotic and mood disorders (including bipolar disorder, rapid-cycling bipolar disorder, and adolescent mania) has been reported in several randomized or open-label studies (17–24).

Based on the need for new treatment options for bipolar depression, the effectiveness of atypical antipsychotics in acute mania, and the emerging evidence for their use in bipolar depression, we evaluated the efficacy and safety of quetiapine compared with placebo in the treatment of depressive episodes in patients with bipolar I or bipolar II disorder.
Method

This double-blind, randomized, fixed-dose, placebo-controlled, parallel-group monotherapy study of quetiapine versus placebo was conducted at 39 centers in the United States between September 2002 and October 2003. After a washout period of at least five half-lives of any prior psychotropic medications, subjects were treated for 8 weeks to evaluate the efficacy, safety, and tolerability of 600 and 300 mg/day of quetiapine and placebo in the treatment of depressive episodes in adult patients with bipolar I or II disorder.

The study was approved by institutional review boards for each site and performed in accordance with the current amendment of the Declaration of Helsinki and the International Conference on Harmonization/Good Clinical Practice guidelines. Written informed consent was obtained from all subjects before participation.

Patient Population

Outpatients ages 18 to 65 years who met DSM-IV criteria for bipolar I or II disorder and were experiencing a major depressive episode were eligible for inclusion in the study. The diagnosis was confirmed with the Structured Clinical Interview for DSM-IV. The patients were required to have a Hamilton Depression Rating Scale 17-item score ≥20 (25), a Hamilton depression scale item 1 score ≥2, and a Young Mania Rating Scale (26) score ≤12 at both the screening and randomization visits. Inclusion criteria were based on the Hamilton depression scale rather than the primary efficacy measure (the Montgomery-Åsberg Depression Rating Scale [27]).

Patients were excluded from the study if they were diagnosed with an axis I disorder other than bipolar disorder that was the primary focus of treatment within 6 months before the screening, if the current episode of depression exceeded 12 months or was less than 4 weeks in duration, or if they had a history of nonresponse to an adequate (6-week) trial of more than two classes of antidepressants during the current episode. Additional exclusion criteria included a diagnosis of substance dependence (DSM-IV) or substance use (except for nicotine) within 12 months before the screening or a clinically significant medical illness. Patients who posed a current serious suicidal or homicidal risk were also excluded. Patients were not permitted to take benzodiazepines during the washout period, and only limited use was permitted during the first 3 weeks after random assignment.

Random assignment was achieved in a non-center-specific manner with an interactive voice-response central randomization service. Random assignment was stratified according to bipolar type (1 or II) to ensure a relative balance in the total number of patients among groups (1:1:1). The patients were randomly assigned to one of three groups: quetiapine, 600 mg/day; quetiapine, 300 mg/day; or placebo.

Study Medication

Quetiapine (600 mg/day or 300 mg/day) or placebo was administered orally, in a single dose, once a day at bedtime. Quetiapine was initiated at 50 mg/day and administered to achieve a target dose of 300 mg/day by day 4 or 600 mg/day by week 1. All packaging of treatments was identical, with placebo and active tablets identical in appearance and number.

Prior and Concomitant Medication

Nonpsychotropic medication, including over-the-counter medications taken before entry into the study could be continued. Zolpidem tartrate (5–10 mg/day at bedtime for insomnia) and lorazepam (1–3 mg/day for severe anxiety) were permitted at the discretion of the investigator and only during the first 3 weeks of treatment but were withheld for 8 hours before psychiatric assessments were conducted. The use of all other psychotropic drugs was prohibited during the study.

Efficacy Evaluations

Clinical assessments were conducted at baseline and weekly from weeks 1 to 8. The primary efficacy variable was the mean change in the Montgomery-Åsberg Depression Rating Scale total score from baseline to week 8 (27).

Additional efficacy evaluations included a change from baseline to each assessment on the Montgomery-Åsberg Depression Rating Scale, the proportion of patients who achieved a protocol-defined response (≥50% reduction from baseline score on the Montgomery-Åsberg Depression Rating Scale), the time to response, the proportion of patients who achieved remission (Montgomery-Åsberg Depression Rating Scale score ≤12), the time to remission, as well as a Montgomery-Åsberg Depression Rating Scale item analysis. The change from baseline to each assessment on the Hamilton depression scale, the Clinical Global Impression (CGI) (28) severity of illness score, and the CGI improvement score were also assessed.

The effect of quetiapine on anxiety symptoms was assessed with the Hamilton Anxiety Rating Scale (29). Mean change from baseline to each assessment and at week 8 in the Hamilton anxiety scale total score was determined.

Quality of sleep was assessed with the Pittsburgh Sleep Quality Index, which measures several dimensions of sleep, including quality, latency, duration, efficiency, use of medication, and daytime dysfunction (30). The 16-item short form of the Quality of Life Enjoyment and Satisfaction Questionnaire was used to measure satisfaction with various areas of daily functioning, such as social relationships, living/housing, physical health, medication, and global satisfaction (31). The Pittsburgh Sleep Quality Index and the Quality of Life Enjoyment and Satisfaction Questionnaire were administered at baseline and at weeks 4 and 8.

Safety and Tolerability Evaluations

Safety and tolerability were evaluated by assessing the incidence and severity of adverse events, as well as withdrawals because of adverse events. Extrapyramidal symptoms were assessed with the Simpson-Angus Rating Scale (32), and akathisia was assessed with the Barnes Rating Scale for Drug-Induced Akathisia (33) at random assignment and at week 8. Measurements of vital signs, including weight and fasting serum glucose levels, were obtained at each study visit. Twelve-lead ECGs, clinical chemistry, and hematology assessments were performed at the screening and at week 8.

The incidence of treatment-emergent mania was evaluated by comparing the percentage of patients in each group who had a total Young Mania Rating Scale score ≥16 on any two consecutive visits or at the final assessment, or an adverse event of mania or hypomania.

Statistical Analyses

Primary and secondary efficacy analyses were performed on the intent-to-treat population, which included all randomly assigned patients who took at least one dose of study medication and had at least one postbaseline efficacy assessment. A last-observation-carried-forward analysis was used to impute missing data for patients who withdrew during the study. All statistical tests were two-tailed. The primary analysis of change from baseline to final assessment in the Montgomery-Åsberg Depression Rating Scale total scores tested the superiority of each dose of quetiapine in the intent-to-treat group (patients with bipolar I or bipolar II disorder) with an analysis of covariance (ANCOVA) with the baseline Montgomery-Åsberg Depression Rating Scale as the covariate and included treatment and diagnosis strata as fixed ef-
the 542 randomly assigned patients, 539 received at least
illustrates the disposition of patients during the study. Of
screening compared with those who were randomly as-
181). There were no significant differences between the
randomly assigned to receive quetiapine, 600 mg/day (N=
with bipolar I (N=360) or bipolar II (N=182) disorder were
screened (N=838) patients who did not pass the
completed study (N=98) Patients selected for random assignment (N=542)
Bipolar I patients (N=360) Bipolar II patients (N=182)
Completed study (N=121) Completed study (N=107)
Completed study (N=110)
Discontinued (N=60) Lost to follow-up (N=12) Adverse event (N=29) Protocol noncompliance (N=10) Withdrawn informed consent (N=5) Lack of efficacy (N=4)
Discontinued (N=82) Lost to follow-up (N=21) Adverse event (N=47) Protocol noncompliance (N=4) Withdrawn informed consent (N=6) Lack of efficacy (N=1) Other (N=3)

TABLE 1. Baseline Demographic Characteristics of
Screened Outpatients With Bipolar I or II Disorder Who Experienced a Major Depressive Episode

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Patients Who Did Not Pass Screening (N=296)</th>
<th>Patients Who Were Randomly Assigned to Treatment (N=539)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Female sex</td>
<td>168</td>
<td>56.8</td>
</tr>
<tr>
<td>Caucasian race</td>
<td>227</td>
<td>76.7</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–39</td>
<td>163</td>
<td>55.1</td>
</tr>
<tr>
<td>40–59</td>
<td>122</td>
<td>41.2</td>
</tr>
<tr>
<td>≥60</td>
<td>10</td>
<td>3.4</td>
</tr>
</tbody>
</table>

\(^a\) Safety population that excluded three patients who did not receive any dose of study medication.

Results

Patients and Disposition

A total of 838 patients were screened, and 542 patients with bipolar I (N=360) or bipolar II (N=182) disorder were randomly assigned to receive quetiapine, 600 mg/day (N=180); quetiapine, 300 mg/day (N=181); or placebo (N=181). There were no significant differences between the baseline characteristics of patients who did not pass the screening compared with those who were randomly assigned (Table 1). The most common reason for the screening failure was failure to meet eligibility criteria. Figure 1 illustrates the disposition of patients during the study. Of the 542 randomly assigned patients, 539 received at least one dose of study medication and were included in the safety population. Of these, 511 had at least one postbaseline assessment and were analyzed for efficacy in the intent-to-treat population.

There were no statistically significant differences between treatment groups with respect to any demographic and baseline disease characteristic (Table 2). The mean age was approximately 37 years, and 58.2% of the patients were women. Mean Montgomery-Åsberg Depression Rating Scale scores at baseline were consistent with moderate to severe depression (34).

There were no statistically significant differences between the quetiapine groups and placebo in the proportion of the patients who completed the study: 54% in the 600 mg/day quetiapine group, 67% in the 300 mg/day quetiapine group, and 59% in the placebo group. The most common reasons for withdrawal were related to adverse events in the quetiapine groups (26.1% and 16.0%) and lack of efficacy in the placebo group (13.3%).

The use of lorazepam and zolpidem (permitted during the first 3 weeks of the study) was generally low across groups. Lorazepam use during the study was 5.6% and 9.5% in the 600 and 300 mg/day quetiapine groups, respectively, compared with 8.3% in the placebo group.

Effects in the model, with adjustment for multiple comparisons. Effect size (improvement of quetiapine over placebo divided by pooled standard deviation) was determined with a mixed-model repeated-measures analysis.

Differences in response rates between treatment and placebo groups and in patients with and without rapid cycling were assessed with a Cochran-Mantel-Haenszel chi-square test across diagnostic strata. Hamilton depression scale, CGI severity and improvement, Young Mania Rating Scale, Hamilton anxiety scale, Pittsburgh Sleep Quality Index, and Quality of Life Enjoyment and Satisfaction Questionnaire scores were tested with ANCOVAs. All secondary analyses were conducted at the nominal significance level of 0.05, with no adjustment for multiple comparisons.

Sample sizes were determined to provide 85% power to detect a difference of 3.6 points on the Montgomery-Åsberg Depression Rating Scale with two-tailed pairwise comparisons between treatment groups and placebo at an alpha level of 0.025 in the intent-to-treat population (patients with bipolar I or bipolar II disorder).

Exploratory analyses were carried out on the bipolar I and II subgroups whose group size was not predetermined to provide power for significance testing. Exploratory analyses were limited to descriptions of the mean changes in primary outcome measure across the three treatment groups, and effect size determinations for the groups taking 600 and 300 mg/day of quetiapine. The repeated measures mixed-effects model included terms for treatment, bipolar diagnosis, treatment-by-bipolar diagnosis, baseline Montgomery-Åsberg Depression Rating Scale total score, visit (week), and treatment-by-visit effects. Several covariance structures were examined, including autoregressive, banded Toeplitz, compound symmetry, and unstructured. The best-fitting covariance structure, the banded Toeplitz, was determined with the Bayesian information criterion.
Zolpidem use during the study was 6.7% and 4.5% in the 600 and 300 mg/day quetiapine groups, respectively, compared with 8.3% in the placebo group.

**Efficacy**

**Montgomery-Åsberg Depression Rating Scale.** Mean baseline Montgomery-Åsberg Depression Rating Scale scores were 30.3 (SD=5.3), 30.4 (SD=5.0), and 30.6 (SD=5.3) in the 600 mg/day, 300 mg/day, and placebo groups, respectively. Quetiapine at a dose of either 600 or 300 mg/day demonstrated significantly greater mean improvement in Montgomery-Åsberg Depression Rating Scale total scores compared with placebo as early as week 1 and at all time points that followed in the intent-to-treat group of patients with bipolar I or II depression (p < 0.001 for both quetiapine doses versus placebo) (Figure 2). The mean change in Montgomery-Åsberg Depression Rating Scale total score from baseline to last assessment was –16.73 in the 600 mg/day group and –16.39 in the 300 mg/day group, compared with –10.26 in the placebo group (p < 0.001 for both quetiapine doses versus placebo) (Table 3, Figure 2). The effect sizes were 0.81 for 600 mg/day and 0.67 for 300 mg/day of quetiapine.

Approximately 58% of the patients treated with either dose of quetiapine were responders at the final assessment, and both doses resulted in significantly higher response rates than placebo (36.1%) (p < 0.001). Notably, the percentage of patients meeting response criteria with 600 mg/day of quetiapine was significantly higher as early as week 1 (24.3%) versus placebo (10.7%) (p < 0.001). In the group taking 300 mg/day of quetiapine, a significantly higher response rate (37.2%) versus placebo (19.5%) was apparent by week 2 (p < 0.001). The median time to response was significantly shorter for both 600 mg/day (22 days) and 300 mg/day (22 days) of quetiapine compared with placebo (36 days) (log-rank χ²=33.1, df=2, p<0.001).

The percentage of patients meeting remission criteria at the final assessment was 52.9% in both the groups taking 600 and 300 mg/day of quetiapine, significantly higher than the placebo rate of 28.4% in each group (p < 0.001). The median time to remission was significantly shorter for

---

### TABLE 2. Baseline Demographic and Clinical Characteristics of Outpatients With Bipolar I or II Disorder Who Experienced a Major Depressive Episode

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Patients Taking Quetiapine 600 mg/day (N=170)</th>
<th>Patients Taking Quetiapine 300 mg/day (N=172)</th>
<th>Patients Taking Placebo (N=169)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>71</td>
<td>41.8</td>
<td>79</td>
</tr>
<tr>
<td>Female</td>
<td>99</td>
<td>58.2</td>
<td>93</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>144</td>
<td>84.7</td>
<td>141</td>
</tr>
<tr>
<td>Black</td>
<td>18</td>
<td>10.6</td>
<td>23</td>
</tr>
<tr>
<td>Hispanic</td>
<td>5</td>
<td>2.9</td>
<td>7</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>1.8</td>
<td>1</td>
</tr>
<tr>
<td>DSM-IV diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar I disorder</td>
<td>114</td>
<td>67.1</td>
<td>116</td>
</tr>
<tr>
<td>Bipolar II disorder</td>
<td>56</td>
<td>32.9</td>
<td>56</td>
</tr>
<tr>
<td>DSM-IV rapid cycling</td>
<td>31</td>
<td>18.2</td>
<td>42</td>
</tr>
<tr>
<td>Mean Age (years)</td>
<td>37.3</td>
<td>11.4</td>
<td>36.6</td>
</tr>
<tr>
<td>Mean Montgomery-Åsberg Depression Rating Scale</td>
<td>30.3</td>
<td>5.3</td>
<td>30.4</td>
</tr>
<tr>
<td>Mean Hamilton Depression Rating Scale</td>
<td>24.7</td>
<td>3.5</td>
<td>24.5</td>
</tr>
<tr>
<td>Mean Hamilton Anxiety Rating Scale</td>
<td>18.7</td>
<td>7.3</td>
<td>18.6</td>
</tr>
<tr>
<td><strong>a</strong> Intent-to-treat analysis.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**FIGURE 2. Least-Squares Mean Change From Baseline in Montgomery-Åsberg Depression Rating Scale Total Score at Each Assessment of Outpatients With Bipolar I or II Disorder Who Experienced a Major Depressive Episode**

---

*a* Intent-to-treat, last-observation-carried-forward analyses. Improvement in Montgomery-Åsberg Depression Rating Scale total score with both doses of quetiapine (600 mg/day and 300 mg/day) was significantly greater than placebo at every assessment (p<0.001).
TABLE 3. Baseline and Mean Change in Efficacy Measures at the Last Assessment of Outpatients With Bipolar I or II Disorder Who Experienced a Major Depressive Episodea

<table>
<thead>
<tr>
<th>Measure and Treatment</th>
<th>Baseline Score</th>
<th>Change in Score at Last Assessment</th>
<th>Analysis (comparison with placebo)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>ANCOVA (df=1)b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p</td>
</tr>
<tr>
<td>Montgomery-Åsberg Depression Rating Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600 mg/day of quetiapine</td>
<td>30.3</td>
<td>5.3</td>
<td>−16.73</td>
</tr>
<tr>
<td>300 mg/day of quetiapine</td>
<td>30.4</td>
<td>5.0</td>
<td>−16.39</td>
</tr>
<tr>
<td>Placebo</td>
<td>30.6</td>
<td>5.3</td>
<td>−10.26</td>
</tr>
<tr>
<td>Hamilton Depression Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600 mg/day of quetiapine</td>
<td>24.7</td>
<td>3.5</td>
<td>−13.84</td>
</tr>
<tr>
<td>300 mg/day of quetiapine</td>
<td>24.5</td>
<td>3.0</td>
<td>−13.38</td>
</tr>
<tr>
<td>Placebo</td>
<td>24.6</td>
<td>3.3</td>
<td>−8.54</td>
</tr>
<tr>
<td>Hamilton Depression Scale item 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600 mg/day of quetiapine</td>
<td>2.9</td>
<td>0.5</td>
<td>−1.68</td>
</tr>
<tr>
<td>300 mg/day of quetiapine</td>
<td>2.9</td>
<td>0.5</td>
<td>−1.65</td>
</tr>
<tr>
<td>Placebo</td>
<td>2.9</td>
<td>0.4</td>
<td>−1.11</td>
</tr>
<tr>
<td>Clinical Global Impression scale Improvement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600 mg/day of quetiapine</td>
<td>4.5</td>
<td>0.6</td>
<td>2.37</td>
</tr>
<tr>
<td>300 mg/day of quetiapine</td>
<td>4.4</td>
<td>0.5</td>
<td>2.27</td>
</tr>
<tr>
<td>Placebo</td>
<td>4.4</td>
<td>0.6</td>
<td>2.97</td>
</tr>
<tr>
<td>Severity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600 mg/day of quetiapine</td>
<td>4.5</td>
<td>0.6</td>
<td>−1.66</td>
</tr>
<tr>
<td>300 mg/day of quetiapine</td>
<td>4.4</td>
<td>0.5</td>
<td>−1.63</td>
</tr>
<tr>
<td>Placebo</td>
<td>4.4</td>
<td>0.6</td>
<td>−0.95</td>
</tr>
<tr>
<td>Hamilton Anxiety Rating Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600 mg/day of quetiapine</td>
<td>18.7</td>
<td>7.3</td>
<td>−8.75</td>
</tr>
<tr>
<td>300 mg/day of quetiapine</td>
<td>18.6</td>
<td>7.3</td>
<td>−8.64</td>
</tr>
<tr>
<td>Placebo</td>
<td>18.9</td>
<td>7.3</td>
<td>−5.54</td>
</tr>
<tr>
<td>Pittsburgh Sleep Quality Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600 mg/day of quetiapine</td>
<td>11.6</td>
<td>4.2</td>
<td>−5.46</td>
</tr>
<tr>
<td>300 mg/day of quetiapine</td>
<td>11.4</td>
<td>3.8</td>
<td>−5.16</td>
</tr>
<tr>
<td>Placebo</td>
<td>11.7</td>
<td>3.8</td>
<td>−2.94</td>
</tr>
<tr>
<td>Quality of Life Enjoyment and Satisfaction Questionnaire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600 mg/day of quetiapine</td>
<td>34.1</td>
<td>8.2</td>
<td>11.71</td>
</tr>
<tr>
<td>300 mg/day of quetiapine</td>
<td>36.1</td>
<td>7.9</td>
<td>10.77</td>
</tr>
<tr>
<td>Placebo</td>
<td>34.2</td>
<td>7.4</td>
<td>6.44</td>
</tr>
</tbody>
</table>

a Intent-to-treat, last-observation-carried-forward analyses.
b Test, treatment contrast within the framework of the ANCOVA, estimated difference (standard error).

Both 600 mg/day (27 days) and 300 mg/day (29 days) of quetiapine compared with placebo (65 days) (log-rank $\chi^2$ = 32.8, df=2, p<0.001).

Nine out of 10 Montgomery-Åsberg Depression Rating Scale items were significantly improved from baseline compared with placebo in the 600 mg/day quetiapine group, as were eight items in the 300 mg/day quetiapine group (p<0.05) (Figure 3). With both doses of quetiapine, these items included the core mood symptoms of apparent sadness, reported sadness, inability to feel, pessimistic thoughts, and suicidal thoughts. The core mood symptoms of apparent sadness, reported sadness, and pessimistic thoughts were significantly improved in both quetiapine groups as early as week 1 compared with placebo (p<0.05). An inability to feel and suicidal thoughts were also significantly improved by week 1 in the group taking 600 mg/day of quetiapine compared with placebo (p<0.05). Both doses of quetiapine were more effective than placebo in reducing suicidal thoughts at the final assessment (p=0.001); the reductions with quetiapine were approximately twice that of placebo.

In the bipolar I subgroup of patients, the mean change in Montgomery-Åsberg Depression Rating Scale total score from baseline to last assessment was −18.05 in the group taking 600 mg/day group of quetiapine and −16.91 in the 300 mg/day group, compared with −9.24 in the placebo group (p<0.001 for both quetiapine doses versus placebo). The effect size in the bipolar I subgroup was 1.09 for those assigned to 600 mg/day and 0.91 for those taking 300 mg/day of quetiapine. In the subgroup of patients with bipolar II disorder, the mean change in Montgomery-Åsberg Depression Rating Scale total scores compared with placebo at most assessments was smaller than in bipolar I patients. Although the change in Montgomery-Åsberg Depression Rating Scale total score from baseline to last assessment occurred with quetiapine treatment regardless of the presence of rapid cycling in the intent-to-treat group (patients with bipolar I or II disorder). The mean change in Montgomery-Åsberg Depression Rat-
A Major Depressive Episode for Outpatients with Bipolar I or II Disorder Experiencing a Rapid Cycling was 1356 Am J Psychiatry 162:7, July 2005

cycling was reduced in the patients with rapid cycling was ~16.6 in the 600 mg/day group and ~15.7 in the 300 mg/day group versus ~10.3 in the placebo group (p<0.001 for both quetiapine doses versus placebo). The mean change in Montgomery-Åsberg Depression Rating Scale total score at week 8 in the patients without rapid cycling was ~18.6 in the 300 mg/day quetiapine group versus ~9.9 in the placebo group (p<0.01 for both quetiapine doses versus placebo). The mean change in Montgomery-Åsberg Depression Rating Scale total score at week 8 in the patients with rapid cycling was ~18.8 in the pooled quetiapine groups (600 or 300 mg/day) versus ~18.9 in the placebo group. In the patients without somnolence/sedation, the mean change in the Montgomery-Åsberg Depression Rating Scale total score was ~19.3 and ~11.7 for in the pooled quetiapine and placebo groups, respectively. The placebo group response was higher in the patients reporting somnolence/sedation, but the results with quetiapine were similar in the patients with or without somnolence/sedation.

**Hamilton depression scale.** Mean baseline Hamilton depression scale scores were 24.7 (SD=3.5), 24.5 (SD=3.0), and 24.6 (SD=3.3) in the 600 mg/day, 300 mg/day, and placebo groups, respectively (Table 2). Quetiapine at a dose of either 600 or 300 mg/day demonstrated significantly greater mean improvements in Hamilton depression scale total scores compared to placebo as early as week 1 and at all time points that followed in the patients with bipolar I or II depression (p<0.001). The mean change from baseline in Hamilton depression scale scores at week 8 was ~13.84, ~13.38, and ~8.54 in the 600 mg/day, 300 mg/day, and placebo groups, respectively (p<0.001 for both quetiapine doses versus placebo). At the end of the study, the effect sizes for the group of patients with bipolar I or II disorder with the Hamilton depression scale was 0.93 for 600 mg/day and 0.74 for 300 mg/day of quetiapine.

Significant improvement in the Hamilton depression scale item 1 (depressed mood) was as early as week 1 (p=0.003) for both quetiapine doses and continued to be statistically superior to placebo at all time points.

**Clinical Global Impression.** Quetiapine-treated patients experienced a statistically significant improvement (p<0.001) on the CGI severity scale as early as week 1 that was sustained to the end of the study for both quetiapine doses versus placebo. At the final assessment, a larger percentage of patients were rated as “normal, not at all ill,” or

### FIGURE 3. Mean Percent Change From Baseline in Individual Montgomery-Åsberg Depression Rating Scale Items for Outpatients with Bipolar I or II Disorder Experiencing a Major Depressive Episode

<table>
<thead>
<tr>
<th>Item</th>
<th>Patients assigned to quetiapine, 600 mg/day (N=170)</th>
<th>Patients assigned to quetiapine, 300 mg/day (N=172)</th>
<th>Patients assigned to placebo (N=169)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apparent sadness</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Reported sadness</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Inner tension</td>
<td></td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>Reduced sleep</td>
<td></td>
<td></td>
<td>b</td>
</tr>
<tr>
<td>Reduced appetite</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concentration difficulties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lassitude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inability to feel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pessimistic thoughts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suicidal thoughts</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Intent-to-treat, last-observation-carried-forward analyses. Nine of 10 and 8 of 10 Montgomery-Åsberg Depression Rating Scale items (including the core mood symptoms of depression [item 1: apparent sadness; item 2: reported sadness; item 8: inability to feel; item 9: pessimistic thoughts; item 10: suicidal thoughts]) were significantly improved from baseline compared to placebo in the groups taking 600 mg/day and 300 mg/day of quetiapine, respectively (p<0.05). Apparent sadness, reported sadness, and pessimistic thoughts were significantly improved in both quetiapine groups as early as week 1 compared with placebo (p<0.05). Both doses of quetiapine were approximately twice as effective as placebo in reducing suicidal thoughts at the final assessment (p<0.01).

b p<0.001 versus placebo.
c p<0.01.
d p<0.05.
quetiapine was significantly greater with both doses compared with placebo. The mean improvement in Pittsburg Sleep Quality Index scores from baseline in patients treated with 600 mg/day of quetiapine and by 10.77 among those treated with 300 mg/day of quetiapine, compared with 6.44 in the placebo group (p<0.001 for both quetiapine doses versus placebo).

### Safety and Tolerability

**Adverse events.** Common adverse events (whether or not considered treatment related) occurred in ≥10% of patients, and withdrawals due to common adverse events are shown in Table 4. The overall rate of study discontinuation due to adverse events was 26.1% (N=47) in the 600 mg/day group, 16.0% (N=29) in the 300 mg/day group, and 8.8% (N=16) in the placebo group (Figure 1). There were no significant differences in the rates of serious adverse events across treatment groups, and none was treatment related: 5.0% (N=9) in the 600 mg/day group and 3.4% (N=6) in the 300 mg/day group compared with 8.9% (N=16) in the placebo group. Two patients attempted suicide (one in each of the active treatment groups), but no suicides or deaths occurred during the study.

The rate of discontinuation due to adverse events in the subgroup of patients with bipolar I disorder was 23.3% (N=28) in the 600 mg/day group, 13.1% (N=16) in the 300 mg/day group, and 11.9% (N=14) in the placebo group. The incidence of serious adverse events in the subgroup of patients with bipolar I disorder was 5.0% (N=6) in the 600 mg/day group, 4.2% (N=5) in the 300 mg/day group, and 11.9% (N=14) in the placebo group.

In the subgroup of patients with bipolar II disorder, the rate of discontinuation due to adverse events was 31.7% (N=19) in the 600 mg/day group, 22.0% (N=13) in the 300 mg/day group, and 3.2% (N=2) in the placebo group. The incidence of serious adverse events in the subgroup of patients with bipolar II disorder was 5.0% (N=3) in the 600 mg/day group, 1.7% (N=1) in the 300 mg/day group, and 3.2% (N=2) in the placebo group.

### Table 4. Incidence and Withdrawals Because of Adverse Events Occurring in at Least 10% of the Patients in Any Group of Outpatients with Bipolar I or II Disorder Who Experienced a Major Depressive Episode

<table>
<thead>
<tr>
<th>Adverse Event</th>
<th>Incidence N (%)</th>
<th>Leading to Withdrawal N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patients Taking 600 mg/day of Quetiapine (N=180)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry mouth</td>
<td>73 (40.6%)</td>
<td>2 (1.1%)</td>
</tr>
<tr>
<td>Sedation</td>
<td>58 (32.2%)</td>
<td>17 (9.4%)</td>
</tr>
<tr>
<td>Somnolence</td>
<td>44 (24.4%)</td>
<td>5 (2.8%)</td>
</tr>
<tr>
<td>Dizziness</td>
<td>41 (22.6%)</td>
<td>6 (3.3%)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>21 (11.7%)</td>
<td>1 (0.6%)</td>
</tr>
<tr>
<td>Constipation</td>
<td>20 (11.1%)</td>
<td>1 (0.6%)</td>
</tr>
<tr>
<td>Headache</td>
<td>18 (10.0%)</td>
<td>1 (0.6%)</td>
</tr>
<tr>
<td>Nausea</td>
<td>6 (8.9%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Upper respiratory tract infection not otherwise specified</td>
<td>13 (7.2%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td><strong>Patients Taking 300 mg/day of Quetiapine (N=179)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry mouth</td>
<td>79 (44.1%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Sedation</td>
<td>53 (29.6%)</td>
<td>10 (5.6%)</td>
</tr>
<tr>
<td>Somnolence</td>
<td>49 (27.4%)</td>
<td>7 (3.9%)</td>
</tr>
<tr>
<td>Dizziness</td>
<td>30 (16.8%)</td>
<td>1 (0.6%)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>16 (8.9%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Constipation</td>
<td>21 (11.7%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Headache</td>
<td>22 (12.3%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Nausea</td>
<td>14 (7.8%)</td>
<td>3 (1.7%)</td>
</tr>
<tr>
<td>Upper respiratory tract infection not otherwise specified</td>
<td>9 (5.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td><strong>Patients Taking Placebo (N=180)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry mouth</td>
<td>14 (7.8%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Sedation</td>
<td>11 (6.1%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Somnolence</td>
<td>15 (8.3%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Dizziness</td>
<td>15 (8.3%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>13 (7.2%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Constipation</td>
<td>8 (4.4%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Headache</td>
<td>36 (20.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Nausea</td>
<td>23 (12.8%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Upper respiratory tract infection not otherwise specified</td>
<td>18 (10.0%)</td>
<td>0 (0.0%)</td>
</tr>
</tbody>
</table>

*Significantly higher than placebo (p<0.05).*
The incidence of treatment-emergent mania was low and not significantly different from placebo at either quetiapine dose: 2.2% with 600 mg/day of quetiapine (Cochran-Mantel-Haenszel, odds ratio=0.57, 95% confidence interval (CI)=0.17–1.91, p=0.35), 3.9% with 300 mg/day of quetiapine (Cochran-Mantel-Haenszel, odds ratio=0.97, 95% CI=0.35–2.68, p=0.95), and 3.9% with placebo.

The mean Simpson-Angus Rating Scale total score decreased in all three groups from baseline to the final assessment by −0.1, −0.2, and −0.3 in the 600 mg/day and 300 mg/day quetiapine groups and the placebo groups, respectively. There was no statistically significant difference in the number of patients with an increase from baseline in Simpson-Angus Rating Scale scores between either of the quetiapine groups and placebo: 15% (logistic regression=0.66, df=3, p<0.08), 9% (logistic regression=0.06, df=3, p=0.89), and 9% in the 600 mg/day and 300 mg/day quetiapine and placebo groups, respectively.

At the last assessment, mean Barnes Rating Scale for Drug-Induced Akathisia scores were low and similar in all groups: 0.3 in the 600 mg/day group, 0.2 in the 300 mg/day group, and 0.1 in the placebo group. There was no statistically significant difference in the number of patients with an increase from baseline in Barnes Rating Scale for Drug-Induced Akathisia score between either of the quetiapine groups and placebo: 12% (logistic regression=0.39, df=3, p=0.31), 9% (logistic regression=0.06, df=3, p=0.89), and 9% in the 600 mg/day and 300 mg/day quetiapine and placebo groups, respectively.

Adverse events considered extrapyramidal symptoms were present in 8.9% of the 600 mg/day group, 6.7% of the 300 mg/day group, and 2.2% of the placebo group; discontinuation rates for extrapyramidal symptoms were 2.8%, 1.1%, and 0.6%, respectively.

**Laboratory Results and Vital Signs**

No clinically relevant differences between groups were seen in the mean change from baseline for any vital signs, ECGs, hematology, or clinical chemistry parameters.

Patients treated with 600 mg/day of quetiapine experienced a mean weight gain of 1.6 kg by the final assessment compared with 1.0 kg in the 300 mg/kg group and 0.2 kg in the placebo group. At the final assessment, 16 patients (9.0%) treated with 600 mg/day of quetiapine, 15 patients (8.5%) treated with 300 mg/day of quetiapine, and three patients (1.7%) who received placebo had a weight gain of ≥7% of their baseline measurement. No patients withdrew from the study because of weight gain.

Mean fasting serum glucose levels at baseline were 86 (SD=12), 87 (SD=13), and 87 (SD=15) mg/dl in the 600 mg/day and 300 mg/day of quetiapine and placebo groups, respectively. By the final assessment, the mean change in fasting serum glucose was 6 mg/dl (SD=17), 3 mg/dl (SD=13), and 4 mg/dl (SD=26) in the 600 mg/day and 300 mg/day of quetiapine and placebo groups, respectively.

**Discussion**

To our knowledge, this is the first randomized, parallel-group, placebo-controlled trial to evaluate the efficacy of quetiapine in bipolar depression. It may also be the first published large-scale, controlled study to assess the efficacy of any pharmacological treatment in a group of patients with bipolar I or II depression, and one of few studies to examine an antidepressant effect in patients with rapid cycling.

Quetiapine monotherapy has significant antidepressant efficacy in a group of patients with bipolar I or II depression based on the primary efficacy analysis (mean change in Montgomery-Åsberg Depression Rating Scale total score from baseline to last assessment). The magnitude of the clinical improvement was substantial and evident from the first assessment (week 1) and at each visit thereafter. The rates of response and remission and the time to response and remission were significantly improved in the quetiapine groups compared with placebo. Compared with placebo, evidence of early and sustained efficacy was observed consistently with both doses of quetiapine and in all secondary efficacy analyses from week 1 onward.

In the Montgomery-Åsberg Depression Rating Scale item analysis, both doses of quetiapine produced a significant and early improvement in all of the core mood symptoms of depression, including objective and reported sadness, anhedonia, and pessimistic thoughts. Notably, both doses of quetiapine were approximately twice as effective as placebo in reducing suicidal ideation. These findings provide support for the conclusion that quetiapine has specific antidepressant properties.

In this study, significant antidepressant efficacy was demonstrated for quetiapine dosed once a day in the evening. This has important clinical relevance because once-daily dosing has been associated with enhanced medication adherence (36). Dosing at bedtime may also offer a means of improving tolerability, particularly regarding somnolence or sedation that are sometimes seen with quetiapine and may help treat the sleep disturbance that often accompanies bipolar depression.

Both doses of quetiapine were associated with improvements in quality of sleep and quality of life and were effective in patients with a recent history of rapid-cycling bipolar disorder. Exploratory analyses suggest that the clinical effect of both doses of quetiapine was greater in patients with bipolar I disorder than those with bipolar II disorder.

The most common side effects of quetiapine included dry mouth, sedation, somnolence, dizziness, and constipation. The most common side effects leading to withdrawal from the study were sedation and somnolence, with most discontinuations occurring within the first week. Of importance, changes in weight observed in all three groups were relatively small and did not result in withdrawal from the study. Quetiapine treatment was not associated with treatment-emergent mania. The long-term safety of quetiapine
is being explored in ongoing bipolar disorder maintenance studies. However, data from patients with schizophrenia does not suggest that unexpected adverse effects during long-term treatment should be expected (37).

Several aspects of the design of this study were innovative. First, the inclusion of patients with bipolar II disorder into a large-scale study of acute bipolar depression was novel and enhanced the generalizability of the findings, particularly since there is a higher incidence of bipolar II disorder than bipolar I disorder. The inclusion of patients with rapid cycling was also innovative and enhanced the generalizability of the findings to this difficult-to-treat subgroup. Second, rather than focusing solely on depressive symptoms, this study included sleep quality and health-related quality-of-life measures. Sleep-quality assessments (both patient- and bed-partner-rated) indicated improvements in functioning in addition to symptom severity, including several dimensions of sleep quality and daytime dysfunction. The quality-of-life scale provided novel information regarding the effect of quetiapine on social relationships, living/housing arrangements, physical health, satisfaction with medication, and global satisfaction. Improvements in these measures provide evidence for improved function and overall quality of life in addition to reduction in the symptoms of the illness.

Moreover, the inclusion of analyses that quantify the magnitude of the clinical effect through effect size determinations gives clinicians useful information. Knowing if a significant difference is caused by a small clinical effect (<0.4), a moderately sized clinical effect (0.40–0.79), or a large clinical effect (>0.79) has the potential of helping the clinician make decisions on how to use a new medication (38). The effect sizes reported in the bipolar I depression study by Tohen et al. (14) were 0.32 with olanzapine monotherapy and 0.68 with olanzapine-fluoxetine combination therapy compared with 1.09 in the bipolar I subgroup with 600 mg/day of quetiapine in this study.

This study had several limitations. First, the number of enrolled patients with bipolar II disorder was not sufficient to draw firm conclusions regarding efficacy in this subgroup. For this reason, post hoc analyses conducted in the bipolar II subgroup included effect size determinations, which are less affected by sample size than significance testing. Second, moderate rates of sedation or somnolence were observed in both quetiapine groups, which might have compromised the integrity of the double-blind design. If this were a significant factor in the assessment of efficacy, the reduction in Montgomery-Åsberg Depression Rating Scale total score in patients experiencing sedation or somnolence would have been greater than those in patients not experiencing these adverse events. However, this was not the case, and the improvements observed on the Montgomery-Åsberg Depression Rating Scale were comparable in patients with or without sedation or somnolence. Third, although the study indicated that the two doses used—chosen because of their efficacy in bipolar mania and other disorders—were effective, guidance on the best dosing for most patients or subgroups of patients should be assessed in future studies.

In conclusion, this large, randomized, double-blind, placebo-controlled study provides the first pivotal data demonstrating that quetiapine monotherapy is efficacious and well tolerated for the acute treatment of bipolar depression in a group of patients with bipolar I or II disorder.

Acknowledgments


Received Aug. 4, 2004; revision received Oct. 28, 2004; accepted Dec. 10, 2004. From the University Hospitals of Cleveland/Case University School of Medicine, Cleveland, Ohio; the Psychopharmacology Research Program, Department of Psychiatry, University of Cincinnati College of Medicine, Cincinnati, Ohio; the Mental Health Care Line and General Clinical Research Center of the Cincinnati Veterans Affairs Medical Center, Cincinnati, Ohio; AstraZeneca, Wilmington, Del.; the Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, Calif.; the Department of Psychiatry and Behavioral Sciences, Duke University, Raleigh, N.C.; and the Department of Psychiatry and Behavioral Medicine, University of South Florida, Tampa. Address correspondence and reprint requests to Dr. Calabrese, 11400 Euclid Ave., Suite 200, Cleveland, OH 44106; joseph.calabrese@uhhs.com (e-mail).

References

22. Support for AstraZeneca Pharmaceuticals (grant 5077US/0049). The authors thank Max Brady, D.Phil., and Aruna Seth, Ph.D., (PAREXEL MMS) for editorial assistance. Financial support for their assistance was provided by AstraZeneca.
Risperidone Treatment of Autistic Disorder: Longer-Term Benefits and Blinded Discontinuation After 6 Months

Objective: Risperidone is effective for short-term treatment of aggression, temper outbursts, and self-injurious behavior in children with autism. Because these behaviors may be chronic, there is a need to establish the efficacy and safety of longer-term treatment with this agent.

Method: The authors conducted a multisite, two-part study of risperidone in children ages 5 to 17 years with autism accompanied by severe tantrums, aggression, and/or self-injurious behavior who showed a positive response in an earlier 8-week trial. Part I consisted of 4-month open-label treatment with risperidone, starting at the established optimal dose; part II was an 8-week randomized, double-blind, placebo-substitution study of risperidone withdrawal. Primary outcome measures were the Aberrant Behavior Checklist irritability subscale and the Clinical Global Impression improvement scale.

Results: Part I included 63 children. The mean risperidone dose was 1.96 mg/day at entry and remained stable over 16 weeks of open treatment. The change on the Aberrant Behavior Checklist irritability subscale was small and clinically insignificant. Reasons for discontinuation of part I included loss of efficacy (N=5) and adverse effects (N=1). The subjects gained an average of 5.1 kg. Part II included 32 patients. The relapse rates were 62.5% for gradual placebo substitution and 12.5% for continued risperidone; this difference was statistically significant.

Conclusions: Risperidone showed persistent efficacy and good tolerability for intermediate-length treatment of children with autism characterized by tantrums, aggression, and/or self-injurious behavior. Discontinuation after 6 months was associated with a rapid return of disruptive and aggressive behavior in most subjects.

Autistic disorder is characterized by impaired social interaction, abnormal language development, and repetitive and restricted patterns of behavior (DSM-IV), and it affects as many as 20 people per 10,000 (1, 2). Children with autism display broad differences in abilities and needs, but accompanying maladaptive behaviors such as self-injurious behavior, aggression, and tantrums are common and frequently severe enough to limit educational and developmental progress. A variety of treatments, including medication, are employed in the management of these maladaptive behaviors. Controlled trials of medication treatment of autism are limited, but evidence provides support for both conventional and atypical antipsychotic agents (3, 4). Among studies of the conventional antipsychotics, well-controlled trials of haloperidol have revealed statistically significant effects, but only modest clinical benefits, in children with autism, and short- and longer-term side effects are of concern (3, 5). Atypical antipsychotics appear to be preferred by clinicians because of the perception that atypicals have a more favorable side effect profile than typical neuroleptics, but few direct comparison data exist. Atypical agents are also of interest because of possible serotonin (5-HT) abnormalities in some individuals with autism and the high affinity of medicines such as risperidone for 5-HT receptors, especially those of the 5-HT2A and 5-HT2C classes (6). Until 2002, only one placebo-controlled study of risperidone in adults with autism (7) and a handful of open-label studies of children with pervasive developmental disorders (8) had been published. In a prior report, we described the short-term efficacy of risperidone over placebo in an 8-week controlled trial for 101 children and adolescents with autistic disorder (4). Risperidone was chosen for study given the greatest preliminary evidence for its efficacy in this population (6, 8). Although the effects of risperidone on aggression, tantrums, and self-injurious behavior were substantial in our short-term study, the question of whether these improvements would endure over time remained unanswered.

In this study, subjects who showed a positive response to risperidone in the short-term trial were enrolled in an additional 4-month open-label trial (total drug exposure=6 months), which was followed by a placebo-controlled discontinuation protocol lasting up to 8 weeks. The aim of the study was threefold: first, to determine if the short-term efficacy of risperidone is maintained over time; second, to determine if the side effect burden of risperidone remains acceptable over an extended treatment period;
TABLE 1. Baseline Characteristics of Children With Autism Who Responded to Risperidone in an 8-Week Trial and Did or Did Not Participate in a 4-Month Open-Label Extension

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participated in Extension Phase (N=63)</th>
<th>Did Not Participate in Extension Phase (N=38)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Demographic profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>8.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male sex</td>
<td>49</td>
<td>77.8</td>
</tr>
<tr>
<td>Tanner stage I or II</td>
<td>55</td>
<td>87.3</td>
</tr>
<tr>
<td>Race or ethnic group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>44</td>
<td>69.8</td>
</tr>
<tr>
<td>Black</td>
<td>6</td>
<td>9.5</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3</td>
<td>4.8</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>4</td>
<td>6.3</td>
</tr>
<tr>
<td>Other (mixed race)</td>
<td>6</td>
<td>9.5</td>
</tr>
<tr>
<td>Parental annual income (dollars)(^b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤20,000</td>
<td>6</td>
<td>9.5</td>
</tr>
<tr>
<td>20,001–40,000</td>
<td>17</td>
<td>27.0</td>
</tr>
<tr>
<td>40,001–60,000</td>
<td>15</td>
<td>23.8</td>
</tr>
<tr>
<td>&gt;60,000</td>
<td>23</td>
<td>36.5</td>
</tr>
<tr>
<td>Parental education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>1</td>
<td>1.6</td>
</tr>
<tr>
<td>High school</td>
<td>13</td>
<td>20.6</td>
</tr>
<tr>
<td>Trade or technical school</td>
<td>4</td>
<td>6.3</td>
</tr>
<tr>
<td>At least some college</td>
<td>35</td>
<td>55.6</td>
</tr>
<tr>
<td>Advanced degree</td>
<td>10</td>
<td>15.9</td>
</tr>
<tr>
<td>Living at home with at least one parent</td>
<td>56</td>
<td>88.9</td>
</tr>
<tr>
<td>Educational placement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular class</td>
<td>4</td>
<td>6.3</td>
</tr>
<tr>
<td>Regular school, special education</td>
<td>47</td>
<td>74.6</td>
</tr>
<tr>
<td>Special school</td>
<td>11</td>
<td>17.5</td>
</tr>
<tr>
<td>Residential school</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>1.6</td>
</tr>
<tr>
<td>Developmental profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average or above</td>
<td>3</td>
<td>4.8</td>
</tr>
<tr>
<td>Borderline</td>
<td>7</td>
<td>11.1</td>
</tr>
<tr>
<td>Mental retardiation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>17</td>
<td>27.0</td>
</tr>
<tr>
<td>Moderate</td>
<td>12</td>
<td>19.0</td>
</tr>
<tr>
<td>Severe</td>
<td>11</td>
<td>17.5</td>
</tr>
<tr>
<td>Profound</td>
<td>7</td>
<td>11.1</td>
</tr>
<tr>
<td>Unable to assess</td>
<td>6</td>
<td>9.5</td>
</tr>
<tr>
<td>Scores on Vineland Adaptive Behavior Scales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>42.6</td>
<td>14.6</td>
</tr>
<tr>
<td>Social(^c)</td>
<td>46.4</td>
<td>11.9</td>
</tr>
<tr>
<td>Daily living</td>
<td>38.5</td>
<td>18.7</td>
</tr>
<tr>
<td>Maladaptive (part 1)</td>
<td>26.4</td>
<td>5.5</td>
</tr>
<tr>
<td>Clinical profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scores on Aberrant Behavior Checklist subscales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritability</td>
<td>26.6</td>
<td>7.5</td>
</tr>
<tr>
<td>Social withdrawal</td>
<td>15.6</td>
<td>7.7</td>
</tr>
<tr>
<td>Stereotypy</td>
<td>10.2</td>
<td>4.3</td>
</tr>
<tr>
<td>Hyperactivity(^d)</td>
<td>33.4</td>
<td>8.5</td>
</tr>
<tr>
<td>Inappropriate speech</td>
<td>6.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Clinical Global Impression severity score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (moderate)</td>
<td>9</td>
<td>14.3</td>
</tr>
<tr>
<td>5 (marked)</td>
<td>36</td>
<td>57.1</td>
</tr>
<tr>
<td>6 (severe)</td>
<td>17</td>
<td>27.0</td>
</tr>
<tr>
<td>7 (extreme)</td>
<td>1</td>
<td>1.6</td>
</tr>
<tr>
<td>Current anticonvulsant treatment</td>
<td>1</td>
<td>1.6</td>
</tr>
<tr>
<td>Past treatment profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medication naive</td>
<td>11</td>
<td>17.7</td>
</tr>
<tr>
<td>Antipsychotic</td>
<td>4</td>
<td>6.5</td>
</tr>
<tr>
<td>Selective serotonin reuptake inhibitor</td>
<td>9</td>
<td>14.5</td>
</tr>
<tr>
<td>Stimulant</td>
<td>16</td>
<td>25.8</td>
</tr>
<tr>
<td>Alpha-2 agonist</td>
<td>11</td>
<td>17.7</td>
</tr>
</tbody>
</table>

\(^a\) The sum of the numbers of subjects may not be the same as the total number because data were missing for some variables.

\(^b\) Nearly significant difference between groups \(\chi^2=6.94, df=3, p<0.08\).

\(^c\) Nearly significant difference between groups \(t=1.74, df=99, p<0.09\).

\(^d\) Significant difference between groups \(t=2.03, df=99, p<0.05\).
and third, to examine the feasibility of risperidone discontinuation after 6 months of treatment.

**Method**

**Setting and Subjects**

This study was an extension of the 8-week double-blind, placebo-controlled trial with parallel groups and the 8-week open-label risperidone treatment offered to placebo nonresponders. These short-term trials were conducted by the Autism Network of the National Institute of Mental Health (NIMH) Research Units on Pediatric Psychopharmacology between June 1999 and April 2001. The five clinical sites included the University of California at Los Angeles, Ohio State University, Indiana University, Yale University, and Kennedy Krieger Institute (Johns Hopkins University). The protocol was approved by the institutional review board at each site and the NIMH Data and Safety Monitoring Board, and written informed consent was obtained from a parent or guardian prior to enrollment. Safety and protocol fidelity were monitored through weekly conference calls involving the principal investigators and study coordinators, through quarterly reviews by the Data and Safety Monitoring Board, and through annual site visits by the coordinating center (Yale University).

This study enrolled 63 of the 101 subjects in the first protocol, all of whom met criteria for autistic disorder as defined in DSM-IV. Other entry criteria (at the time of enrollment in the initial double-blind phase) included 1) age of 5–17 years, 2) significant tantrums, aggression, self-injurious behavior, and/or agitation, 3) absence of significant medical problems and any other psychiatric disorder requiring drug therapy (e.g., bipolar disorder, psychosis), 4) weight of at least 15 kg, and 5) mental age of at least 18 months. No concomitant treatment with psychotropic medication was allowed during any phase of the study, except anticonvulsant treatment for seizure control if the child had been taking a stable dose for 4 weeks and had been free of seizures for 6 months (see reference 9 for a detailed description).

**Protocol Schedule and Design**

At the final visit of the initial 8-week controlled study, or the 8-week open-label risperidone treatment for placebo nonresponders, subjects deemed responders were offered enrollment in the extension protocol. Written informed consent for continued participation was obtained from the parents, and for 10 subjects deemed to have the capacity to consent, assent was obtained. The subjects were then seen every 4 weeks for 4 months for assessment of efficacy, safety, and possible need for dose adjustments. At the end of these 4 months of open-label treatment, subjects who continued to show a positive response entered the discontinuation phase. In this phase, the subjects were randomly assigned again, this time either to continued risperidone at the stable dosage for the third week, and placebo only by 25% per week. Thus, the dose was 75% of the week 16 dose for 1 week, followed by 50% of the week 16 dose for the second week, 25% of the week 16 dose for the third week, and placebo only by the fourth week. All subjects were seen weekly for a total of 8 weeks in the discontinuation phase.

**Baseline Assessment and Measurement of Outcome**

At screening for participation in the initial 8-week treatment study, autistic disorder was ascertained through a history and examination by a research team and was corroborated by the Autism Diagnostic Interview—Revised (10), administered by a clinician trained to reliability (11). The screening measures also included intelligence testing, the Vineland Adaptive Behavior Scales (12, 13), routine laboratory tests, an electrocardiogram, measurements of height, weight, and vital signs, a medical history, and a physical examination. All subjects were also required to manifest clinically significant problems consisting of aggression, tantrums, and/or self-injury as defined by a score of 18 or higher on the irritability subscale of the Aberrant Behavior Checklist—Community version (14, 15) rated by the parent (or primary caretaker) and confirmed by clinician review. Scores on this 15-item subscale range from 0 to 45, with higher scores indicating greater pathology. In addition, the subjects were required to receive a score of moderate or higher on the Clinical Global Impression (CGI) severity scale (16, pp. 218–222) from the examining clinician.

Eligible subjects were randomly assigned to receive risperidone or placebo for 8 weeks; details are provided elsewhere (4). The primary outcome measures were the parent-rated irritability subscale of the Aberrant Behavior Checklist and the clinician-rated CGI improvement scale. Subjects showing a 25% reduction on the irritability subscale and a rating of much improved or very much improved on the CGI improvement scale by the blinded clinical evaluator were classified as positive responders and were eligible for the 4-month open-label extension phase. At the end of the 4-month extension, responders were eligible for the randomized discontinuation phase. In the discontinuation phase, relapse was defined as a 25% increase in the score on the parent-rated Aberrant Behavior Checklist irritability subscale and a CGI improvement rating of much worse or very much worse, compared to the predischortion baseline, for at least 2 consecutive weeks, as assessed by a blinded clinician.

Secondary parent-rated outcome measures included the four additional subscales of the Aberrant Behavior Checklist: lethargy/social withdrawal, stereotypic behavior, hyperactivity, and inappropriate speech. At baseline, the parents were interviewed to identify the child's target symptoms and to rate compulsive behaviors on the Children's Yale-Brown Obsessive Compulsive Scale (see references 4, 13, and 17 for details on study assessment).

**Medication Dosing**

The medication schedule in the initial 8-week trial was based on the child's weight and clinical response. The maximum risperidone dose was 2.5 mg/day for children between 15 and 45 kg and 3.5 mg/day for children weighing above 45 kg. Dose reductions to manage side effects were allowed at any time. During the 4-month open phase, clinicians were allowed to adjust the total daily dose according to response and/or adverse events, with the total daily dose increased up to a maximum of 3.5 mg/day for children weighing 15–45 kg and up to 4.5 mg/day for children above 45 kg.

**Safety Monitoring**

The routine laboratory tests, electrocardiogram, and physical examination were repeated at entry into the extension phase and prior to the discontinuation. Weight and vital signs were assessed at each visit. At each visit, the primary clinician asked the parents about the child's health complaints, intercurrent illness, and concomitant medications, and the clinician administered a 32-item checklist concerning the child's energy level, muscle stiffness, motor restlessness, bowel and bladder habits, sleep, and appetite. Neurological side effects were further assessed at each visit with the Simpson-Angus Rating Scale (18) and the Abnormal Involuntary Movements Scale (AIMS) (16, pp. 534–537). Adverse events indicated by any of these methods were documented according to severity, duration, attribution, outcome, and action taken.
TABLE 2. Scores on Subscales of the Aberrant Behavior Checklist for 63 Children With Autism Who Responded to Risperidone in an 8-Week Trial and Participated in a 4-Month Open-Label Extension

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Week 0a</th>
<th>Endpointb</th>
<th>Analysis (df=1, 58)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Irritability</td>
<td>9.5</td>
<td>6.8</td>
<td>11.7</td>
</tr>
<tr>
<td>Lethargy/social</td>
<td>7.3</td>
<td>5.4</td>
<td>6.8</td>
</tr>
<tr>
<td>withdrawal</td>
<td>4.9</td>
<td>4.3</td>
<td>5.8</td>
</tr>
<tr>
<td>Stereotypy</td>
<td>15.1</td>
<td>10.0</td>
<td>15.8</td>
</tr>
<tr>
<td>Hyperactivity</td>
<td>3.4</td>
<td>3.6</td>
<td>3.4</td>
</tr>
</tbody>
</table>

a Week 0 corresponds to the end of the initial 8 weeks of medication exposure.
b For patients who discontinued treatment, the last observation was carried forward.

Results

Baseline Characteristics

Of the original 101 subjects (82 boys and 19 girls) enrolled in the short-term trial, 63 subjects (mean age=8.6 years, SD=2.8) showed a positive response to risperidone and consented to participate in the extension protocol. Of these 63 subjects, 30 were classified as responders during the double-blind trial and 30 were so classified in the 8-week open-label risperidone trial for placebo nonresponders. There were no differences in scores on the irritability subscale of the Aberrant Behavior Checklist or in the distribution of CGI severity scores at the baseline of the extension phase between subjects who entered directly from the double-blind study and those who entered from the 8-week open-label trial for placebo nonresponders. Therefore, the two groups were combined in the efficacy analysis. During the data analysis, three subjects were noted to have entered the extension phase without fully meeting the response criteria. Inclusion of these three subjects did not alter the results, and thus they were included in all analyses. Demographic characteristics of the entire 63 subjects in the extension study are shown in Table 1. When compared to the 38 subjects who did not participate in the extension phase, the extension group showed few differences. Of a total of 24 comparisons, only one variable, the baseline score on the Aberrant Behavior Checklist hyperactivity subscale, differed significantly, being slightly higher at baseline in the subjects continuing in the extension study.

4-Month Open-Label Treatment: Outcomes

A total of 51 subjects (81.0%) completed the 16-week open-label treatment period. Of the 12 dropouts from the extension study, five subjects were withdrawn because of loss of efficacy, one was withdrawn because of noncompliance with the protocol, one withdrew because of an adverse event (constipation), one withdrew consent (no longer interested in medication treatment), and four were lost to follow-up. The mixed effects regression model for all 63 subjects in the intent-to-treat analysis revealed a significant time effect for the irritability subscale of the Aberrant Behavior Checklist (Table 2, Figure 1). However, the mean irritability ratings showed only a minor increase across the 16-week extension phase, from a mean score of 9.5 (SD=6.8) at the start to 10.8 (SD=7.1) at week 16 (Figure 1), in contrast to the mean pretreatment irritability score...
of 26.6 (SD=7). This 1-point increase lacks clinical significance. Indeed, after 16 weeks of treatment the mean score still showed a 59% reduction from the mean rating at the initiation of risperidone treatment 6 months before, a percent reduction identical to that seen after the initial 8-week double-blind study and yielding an effect size of >1.0. Analysis of the CGI improvement ratings at endpoint showed that 82.5% of the subjects (N=52) continued to be rated as much improved or very much improved (1 or 2 on the CGI improvement scale) (Table 3). By contrast, only 7.9% were rated as worse or much worse compared to baseline.

The risperidone doses were also examined to determine whether a dose increase is required to ensure stability of baseline. The mean risperidone doses were 1.96, 1.80, 1.87, 2.10, and 2.08 mg/day for weeks 0, 4, 8, 12, and 16, respectively, representing a modest 6% increase over baseline.

### 4-Month Open-Label Treatment: Adverse Events

Adverse events, especially mild to moderate increased appetite, tiredness, and/or drowsiness, were common (Table 4). Only one subject withdrew because of an adverse event (constipation). Although the parents of six subjects (9.5%) reported “abnormal movements,” no dyskinesias were identified on repeated examination with the AIMS and Simpson-Angus scale by a physician. One subject with preexisting obesity had galactorrhea according to the mother’s report, but this was not observed on examination. Compared to their weight at the initiation of risperidone treatment, the subjects showed a 6-month weight increase of 5.1 kg (SD=3.6) (paired t=7.46, df=31, p<0.001), which was significantly greater (p<0.001) than the amount expected on the basis of available developmental norms. This finding has been reviewed in more detail in a separate report (20).

### Discontinuation Phase

A total of 38 subjects were enrolled in the discontinuation phase and the remainder of the 63 subjects, aside from the discontinuation subjects’ again showing higher mean baseline scores on the Aberrant Behavior Checklist hyperactivity factor (mean=34.4, SD=8.7, versus mean=30.6, SD=9.0; t=2.08, df=99, p=0.04). Also, subjects in the discontinuation phase showed higher mean baseline scores on the Aberrant Behavior Checklist irritability factor than the remaining 63 subjects (mean=27.6, SD=6.1, versus mean 24.8, SD=7.7, t=2.02, df=99, p=0.05). As planned, the NIMH Data and Safety Monitoring Board independently reviewed two interim efficacy analyses, first after the initial 16 subjects completed this phase and again after the first 32 subjects finished. The relapse rate after 32 subjects completed this phase was significantly higher in the placebo-treated group (62.5%, N=10) than in the group receiving continued risperidone (12.5%, N=2) (Yates’s corrected χ²=6.53, df=1, p=0.01). The median survival time, i.e., time to relapse, was 34 days for the placebo-treated subjects versus 57 days for those continuing to take risperidone (Figure 2). On the basis of the results of this planned interim analysis, the NIMH Data and Safety Monitoring Board ruled that the discontinuation phase be stopped immediately. The four subjects who were still in this phase of the protocol exited the study and were treated clinically. Exploratory post hoc analyses were performed in an effort to identify any clinical predictors of relapse. Age, initial Aberrant Behavior Checklist irritability score, and IQ all failed to differ significantly (p>0.30) between the relapsing and nonrelapsing subjects.

### Discussion

Data from this study suggest that risperidone is a well-tolerated and effective treatment for up to 6 months for children with autism complicated by tantrums, aggression, and self-injury. As measured by the primary indices of response, the CGI improvement scale and the Aberrant Behavior Checklist irritability subscale, improvements associated with risperidone administration were maintained in over 80% of the subjects, with very good tolerability. Furthermore, gradual substitution of placebo for risperidone was associated with a greater return of aggression, temper outbursts, and self-injurious behavior than in the subjects who continued to receive risperidone in the discontinuation phase. Taken together, these data...
RISPERIDONE AND AUTISM

strongly suggest that risperidone is an efficacious treatment for short- and intermediate-term management of serious behavioral problems in children with autism.

As in our study of short-term risperidone efficacy (4), risperidone was also associated with continued maintenance of improvements on the Aberrant Behavior Checklist subscales for hyperactivity, stereotypic behavior, and lethargy/social withdrawal. In addition, the mean reduction from baseline in the irritability subscale scores of 59% at the last observation in the extension phase was nearly identical to that observed in the risperidone group at the completion of the 8-week double-blind efficacy trial. These data expand the limited published database on extended treatment of autistic disorder with medication. Heretofore, the largest extended study of neuroleptic treatment for autistic disorder was the examination of the effects of haloperidol over a 6-month period. In that study (5), 60 children treated with a mean dose of 1.23 mg/day of haloperidol were classified as "good responders." After 6 months of treatment, 33 (63%) of 52 subjects were rated as "much improved." Autism factor ratings derived from the Children’s Psychiatric Rating Scale obtained on entry and at 6 months showed an average decline of 23% from baseline. Upon abrupt withdrawal of haloperidol, only 14% of the children were rated as much worse and 50% showed at least mild deterioration.

In an open-label prospective study of risperidone for 11 children with autistic disorder, Zuddas et al. (21) noted that 10 of the 11 showed enduring improvement over the 12-month observation period. Another open trial (22) of risperidone enrolled 22 subjects with autistic disorder. Of these, 15 subjects showed significant improvement at a mean dose of 1.2 mg/day for up to 7 months of treatment. Additional published observations include positive responses maintained over a 4-month period in responders (23) and improvement maintained during a 2-year treatment history in two adults with autistic disorder (24). It should be noted that few earlier studies defined specific entry criteria for aggression or disruptive behavior, but, taken together, the available accounts suggest that risperidone’s benefit for aggression, severe tantrums, and self-injury accompanying autistic disorder persisted over these intermediate treatment periods (6–12 months), although additional long-term treatment data focusing on managing severe and challenging behaviors are clearly needed.

A key clinical question concerns the length of continued treatment with risperidone for this clinical indication. On one hand, the return of aggression, tantrums, and agitation was five times as great in the placebo-substitution group as in the subjects who continued to take risperidone. On the other hand, 37.5% (six of 16) of the children in the placebo-substitution group did not relapse during the 2-month discontinuation period, demonstrating some

<table>
<thead>
<tr>
<th>Adverse Event During 4-Month Extension</th>
<th>Number of Subjects With Moderate or Severe Event</th>
<th>Percent of Subjects With Moderate or Severe Event</th>
<th>Subjects With Events Related to Study Drug</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moderate</td>
<td>Severe</td>
<td></td>
</tr>
<tr>
<td>Nasal congestion</td>
<td>7</td>
<td>0</td>
<td>11.1</td>
</tr>
<tr>
<td>Appetite increase</td>
<td>5</td>
<td>0</td>
<td>7.9</td>
</tr>
<tr>
<td>Coughing</td>
<td>4</td>
<td>0</td>
<td>6.3</td>
</tr>
<tr>
<td>Anxiety</td>
<td>3</td>
<td>0</td>
<td>4.8</td>
</tr>
<tr>
<td>Difficulty falling asleep</td>
<td>3</td>
<td>0</td>
<td>4.8</td>
</tr>
<tr>
<td>Fever</td>
<td>3</td>
<td>0</td>
<td>4.8</td>
</tr>
<tr>
<td>Skin irritation</td>
<td>3</td>
<td>0</td>
<td>4.8</td>
</tr>
<tr>
<td>Vomiting</td>
<td>3</td>
<td>0</td>
<td>4.8</td>
</tr>
<tr>
<td>Accidental injury</td>
<td>2</td>
<td>0</td>
<td>3.2</td>
</tr>
<tr>
<td>Constipation</td>
<td>2</td>
<td>0</td>
<td>3.2</td>
</tr>
<tr>
<td>Drowsiness/sedation</td>
<td>2</td>
<td>0</td>
<td>3.2</td>
</tr>
<tr>
<td>Enuresis</td>
<td>1</td>
<td>0</td>
<td>3.2</td>
</tr>
<tr>
<td>Headache</td>
<td>2</td>
<td>0</td>
<td>3.2</td>
</tr>
<tr>
<td>Hypersalivation</td>
<td>2</td>
<td>0</td>
<td>3.2</td>
</tr>
<tr>
<td>Sore throat</td>
<td>2</td>
<td>0</td>
<td>3.2</td>
</tr>
<tr>
<td>Depression</td>
<td>0</td>
<td>1</td>
<td>1.6</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Difficulty urinating</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Dyskinesia</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Earache</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Restlessness/agitation</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Sinus condition</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Stomach/abdominal discomfort</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Tiredness/fatigue</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Tremor</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Weight gain</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Muscle rigidity</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>0</td>
<td>1.6</td>
</tr>
</tbody>
</table>

a Mild events are not reported. The most severe event is reported if the child reported the event more than once. One subject reportedly experienced galactorrhea (see text).
degree of variability in outcome. It is conceivable that more gradual drug tapering may have moderated the observed relapse in the risperidone group or that dose reduction, rather than complete medication discontinuation, may have been sufficient to maintain treatment gains. It is also conceivable that concomitant behavioral treatment reinforcing replacement behaviors could protect against relapse following risperidone withdrawal. Nevertheless, the high rate of relapse observed in our study suggests caution when withdrawal of effective treatment for these target symptoms is considered. At a minimum, clinicians should document a clear period of stable functioning before initiating medication taper (25) and ensure that appropriate psychosocial supports are in place to minimize relapse risk. No factors predicting relapse were identifiable given the modest number of subjects in the discontinuation phase of this study, but they certainly form an important future research question. Longer-term follow-up information on outcome may help clinicians make decisions about maintenance treatment. To this end, we are in the process of obtaining 18-month safety data for our study group.

Although adverse events associated with risperidone exposure were observed to be common in this study, most were not deemed clinically significant, and many were not attributed to risperidone. It is important that only one subject withdrew from the 4-month open-label treatment because of an adverse event, and no cases of dyskinesia were observed during 6 months of treatment or upon withdrawal. The absence of dyskinesia in this study is noteworthy given the report by Campbell and colleagues (26) that 30% of 118 subjects showed dyskinesia following withdrawal of haloperidol after a similar duration of treatment, although the more gradual taper used in this study may be responsible for the finding. Weight gain associated with treatment was significant and in some, but not all, cases posed a concern (20), especially since antipsychotic-related weight gain has been associated with diabetes and other adverse outcomes. Longer-term observations to determine the clinical significance of weight gain as well as other adverse events are needed to evaluate the risk-benefit ratio for risperidone treatment in children with autistic disorder.

There were several limitations in this investigation. First, although the data collection period in this phase of the study spanned up to a maximum total of 8 months of risperidone exposure, this does not completely mimic clinical practice, as many patients receive treatments like risperidone for longer time periods. Also, there was no systematic effort to reduce or constrain individual daily doses over time, leaving some uncertainty about the lowest possible long-term dose. This limitation notwithstanding, the relatively low mean dose (approximately 2.0 mg/day) of risperidone used in this study was effective in managing the specific target symptoms of aggression, agitation, and self-injury. This mean dose was remarkably similar to the doses in several other studies (16, 22, 27, 28). The possibility of solidifying these gains or even extending the benefit of risperidone treatment by combining it with behavior therapy was not explored in this investigation but is an important research question for future studies. Finally, the safety observations of the study were limited to a maximum of 8 months of risperidone exposure in a relatively small study group. Thus, our data may be insufficient to estimate precisely the long-term risks of risperidone in children.

In summary, intermediate-length treatment with risperidone appeared to be associated with the maintenance of reductions in aggression, agitation, self-injury, and irritability from short-term treatment. There was little evidence of accommodation to the effects of risperidone, and the medication appeared to be well tolerated in a group of children and adolescents with autistic disorder. Additional studies on predictors of stable improvement, longer-term safety, and approaches combining other interventions are of interest.

Acknowledgments

The following are the authors of this report listed by role and study site: Ohio State University, principal investigator Michael G. Aman, Ph.D., co-investigators L. Eugene Arnold, M.Ed., M.D.,
Ronald Lindsay, M.D., Patricia Nash, M.D., Jill Hollway, B.A.; University of California at Los Angeles, principal investigator James T. McCracken, M.D., co-investigators Bhavik Shah, M.D., James McGough, M.D., Pegeen Cronin, Ph.D., Lisa Lee, B.A.; Indiana University, principal investigator Christopher J. McCougle, M.D., co-investigators David Posey, M.D., Naomi Swiezy, Ph.D., Arlene Kohn, B.A.; Yale University, principal investigator, Lawrence Scahill, M.S.N., Ph.D., co-investigators Andres Martin, M.D., Kathleen Koenig, M.S.N., Fred Volkmar, M.D., Deirdre Carroll, M.S.N., Allison Lancer, B.S.; Kennedy Krieger Institute, principal investigator Elaine Tierney, M.D., co-investigators Jaswinder Ghuman, M.D., Nilda Gonzalez, M.D., Marco Grados, M.D.; National Institute of Mental Health, principal investigator Benedetto Vitiello, M.D., co-investigator Louise Ritz, M.B.A.; Columbia University, statistician Mark Davies, M.P.H.; Nathan Kline Institute, data management James Robinson, M.Ed., Don McMahon, M.S.

Received Oct. 23, 2003; revision received Sept. 22, 2004; accepted Oct. 1, 2004. From the Research Units on Pediatric Psychopharmacology Autism Network (individual contributors are listed in Acknowledgments). Address correspondence and reprint requests to Dr. McCracken, UCLA Neuropsychiatric Institute, 760 Westwood Plaza, Los Angeles, CA 90024-1759; jmccracken@mednet.ucla.edu (e-mail).

Supported by NIMH contracts N01 MH-70010 (principal investigator: Dr. McCracken; N01 MH-70009; Dr. Scahill), N01 MH-70001 (principal investigator: Dr. McCougle), and N01 MH-80011 (principal investigator: Dr. Aman); by NIH Division of Research Resources General Clinical Research Center grants M01 RR-00750 (to Indiana University), M01 RR-00052 (to Johns Hopkins University), M01 RR-00034 (to Ohio State University), and M01 RR-06022 (to Yale University); by NIMH grant MH 01805 (to Dr. McCracken); and by funding from the Koscak Foundation (to Dr. Scahill). Study medications were donated by Janssen Pharmaceutica.

The authors thank the following members of the Autism Network of the NIMH Research Units on Pediatric Psychopharmacology Scientific Advisory Board: C.T. Gordon, M.D., Joseph DeVeaughe-Geis, M.D., Henrietta Leonard, M.D., Richard Shader, M.D., and Susan Swedo, M.D., for their contributions; the NIMH Data and Safety Monitoring Board and Phillip Watson, M.D., for their comments and guidance; and Erin Kustan and Caileng Fu for assistance in preparing this report.

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.

References

Generalizability of Antidepressant Efficacy Trials: Differences Between Depressed Psychiatric Outpatients Who Would or Would Not Qualify for an Efficacy Trial

Mark Zimmerman, M.D.
Iwona Chelminski, Ph.D.
Michael A. Posternak, M.D.

Objective: In recent years the generalizability of antidepressant efficacy trials has been questioned. Central to the question of generalizability is whether there are differences in clinical, demographic, and psychosocial characteristics between patients who would qualify for an antidepressant efficacy trial and those who would not qualify.

Method: The authors compared three groups: 123 depressed patients who would qualify for an antidepressant efficacy trial, 289 whose symptom severity was too mild to qualify for an antidepressant efficacy trial, and 187 who would be excluded because they were suicidal or had a comorbid anxiety or substance use disorder.

Results: Compared with patients who would qualify for an antidepressant efficacy trial, patients who would be excluded because of comorbidity or suicidality were a more chronically ill group with more previous episodes, greater psychosocial impairment, and more personality pathology.

Conclusions: These findings support further caution in generalizing the results from antidepressant efficacy trials to clinical populations.

During the past few years there have been increasing concerns about the generalizability of the results of tightly controlled efficacy trials to real-world clinical practice (1). We recently applied typically used exclusion criteria to a large group of depressed outpatients and found that most depressed patients treated in routine clinical practice would not have qualified for an antidepressant efficacy trial (2). In the present report from the Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project, we turn to the question of the comparability of the patients who would or would not qualify for an antidepressant efficacy trial. Central to the question of generalizability is whether there are important differences in characteristics that are not the basis for exclusion between patients who would or would not qualify for an antidepressant efficacy trial.

Method

The study group included 599 psychiatric outpatients 18 years old or older whose principal diagnosis was DSM-IV nonbipolar, nonpsychotic major depressive disorder. The majority were women (65.4% [N=392]); their mean age was 39.2 years (SD=12.4).

On presentation for outpatient treatment, all patients were interviewed by a trained diagnostic rater who administered the Structured Clinical Interview for DSM-IV supplementing with questions from the Schedule for Affective Disorders and Schizophrenia (SADS) (3) assessing the severity of symptoms during the week preceding the evaluation. An extracted 21-item Hamilton Depression Rating Scale score (4) was derived from the SADS ratings. The interview also included items from the SADS on best level of social functioning during the past 5 years, social functioning during adolescence, and the amount of time employed during the past 5 years. For patients with major depressive disorder we determined the number of episodes, duration of the current episode, age at onset, lifetime history of suicide attempts, and number of hospitalizations. Personality disorders were assessed in 391 of the patients with the Structured Interview for DSM-IV Personality (5). The Rhode Island Hospital Institutional Review Board approved the research protocol, and all patients provided informed consent. The training of the raters and the reliability of the diagnostic assessments in the MIDAS project have been described in detail elsewhere (6).

The present report is based on nonpsychotic patients because treatment recommendations differ for psychotic depression. The four criteria used to define the excluded group are those used in the majority of antidepressant efficacy trials (minimum severity on the Hamilton depression scale, substantial suicide risk, recent DSM-IV diagnosis of alcohol or drug abuse or dependence, and presence of a comorbid axis I disorder) (7). The comorbid disorders used as the basis for exclusion were current posttraumatic stress disorder (PTSD), panic disorder, generalized anxiety disorder, or obsessive-compulsive disorder (OCD).

We decided a priori to subdivide the excluded group into those who would be excluded because of low symptom severity and those who would be excluded because of comorbidity or suicidality, because these variables have different prognostic implications (8). We conducted two planned two-group comparisons of the included patients with each of the excluded groups. A comparison of the two excluded groups was not conducted because it was peripheral to our central question of whether patients who are included or excluded from an antidepressant efficacy trial differ.

Results

The mean extracted Hamilton depression scale score for all 599 patients was 20.3 (SD=5.9). Forty-eight percent of the patients (N=289) scored below 20, the most commonly used severity threshold for inclusion in an efficacy trial (9). Twelve percent of the patients (N=74) had evidence of a current drug or alcohol use disorder. Thirty-nine patients (6.5%) were rated 4 or higher on the SADS suicidal ide-
ation item, indicating the presence of frequent suicidal thoughts with a plan. Forty-one percent (N=246) had current PTSD, generalized anxiety disorder, OCD, or panic disorder. Based on all four criteria, 79.5% (N=476) of the 599 patients would be excluded from an antidepressant efficacy trial.

We compared demographic, psychosocial, and clinical characteristics of the 123 patients who would qualify for an antidepressant efficacy trial with those of the 289 patients whose symptom severity was too mild to qualify for an antidepressant efficacy trial and the 187 patients who would be excluded because they were suicidal or had a comorbid anxiety or substance use disorder (Table 1). The only difference between the included patients and the group excluded for low symptom severity was on the Global Assessment of Functioning Scale (DSM-IV, p. 32), which was expected because symptom severity is one of the components of this scale.

### TABLE 1. Demographic and Psychosocial Characteristics of Depressed Outpatients Who Would Be Included or Excluded From Antidepressant Efficacy Trials

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Included (N=123)</th>
<th>Low Severity (N=289)</th>
<th>Suicidal Ideation or Comorbidity (N=187)</th>
<th>Excluded Analysis</th>
<th>Included Versus Excluded for Low Severity</th>
<th>Included Versus Excluded for Suicidal Ideation or Comorbidity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>χ²</td>
<td>p</td>
<td>χ²</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>42 34.1</td>
<td>92 31.8</td>
<td>73 39.0</td>
<td>0.2</td>
<td>n.s.</td>
<td>0.8 n.s.</td>
</tr>
<tr>
<td>Female</td>
<td>81 65.9</td>
<td>197 68.2</td>
<td>114 61.0</td>
<td>1.0</td>
<td>n.s.</td>
<td>0.7 n.s.</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td>3.6</td>
<td>n.s.</td>
<td>4.2 n.s.</td>
</tr>
<tr>
<td>White</td>
<td>108 87.8</td>
<td>263 91.0</td>
<td>158 84.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonwhite</td>
<td>15 12.2</td>
<td>26 9.0</td>
<td>29 15.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>55 44.7</td>
<td>135 46.7</td>
<td>66 35.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with someone as if married</td>
<td>5 4.1</td>
<td>10 3.5</td>
<td>14 7.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>2 1.6</td>
<td>5 1.7</td>
<td>3 1.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separated</td>
<td>11 8.9</td>
<td>20 6.9</td>
<td>18 9.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divorced</td>
<td>22 17.9</td>
<td>36 12.5</td>
<td>32 17.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>28 22.8</td>
<td>83 28.7</td>
<td>54 28.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td>5.6</td>
<td>n.s.</td>
<td>5.6 n.s.</td>
</tr>
<tr>
<td>Less than high school education</td>
<td>17 13.8</td>
<td>20 6.9</td>
<td>25 13.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school graduation</td>
<td>22 17.9</td>
<td>65 22.5</td>
<td>52 27.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>52 42.3</td>
<td>123 42.6</td>
<td>77 41.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College graduation</td>
<td>32 26.0</td>
<td>81 28.0</td>
<td>33 17.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past social functioning (12–18 years)a</td>
<td></td>
<td></td>
<td></td>
<td>0.2</td>
<td>n.s.</td>
<td>5.5 &lt;0.05</td>
</tr>
<tr>
<td>Good or excellent</td>
<td>92 74.8</td>
<td>209 72.3</td>
<td>116 62.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair or worse</td>
<td>31 25.2</td>
<td>79 27.3</td>
<td>71 38.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current social functioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(past 5 years)b</td>
<td></td>
<td></td>
<td></td>
<td>0.2</td>
<td>n.s.</td>
<td>1.7 n.s.</td>
</tr>
<tr>
<td>Good or excellent</td>
<td>82 66.7</td>
<td>199 68.9</td>
<td>111 59.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair or worse</td>
<td>41 33.3</td>
<td>89 30.8</td>
<td>76 40.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time out of work (past 5 years)b</td>
<td></td>
<td></td>
<td></td>
<td>0.2</td>
<td>n.s.</td>
<td>7.9 &lt;0.01</td>
</tr>
<tr>
<td>Virtually no time</td>
<td>41 39.4</td>
<td>111 41.9</td>
<td>41 23.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substantial absence</td>
<td>63 60.6</td>
<td>154 58.1</td>
<td>133 76.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Episode duration of 2 years or more</td>
<td>37 30.1</td>
<td>113 39.1</td>
<td>80 42.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More than two episodes</td>
<td>36 29.3</td>
<td>89 30.8</td>
<td>78 41.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least one hospitalization</td>
<td>24 19.5</td>
<td>40 13.8</td>
<td>45 24.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least one suicide attempt</td>
<td>22 17.9</td>
<td>48 16.6</td>
<td>59 31.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personality disorder diagnosisc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any personality disorder</td>
<td>19 31.7</td>
<td>65 31.2</td>
<td>70 56.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any cluster A personality disorder</td>
<td>5 8.3</td>
<td>6 2.9</td>
<td>13 10.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any cluster B personality disorder</td>
<td>5 8.3</td>
<td>11 5.3</td>
<td>35 28.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any cluster C personality disorder</td>
<td>14 23.3</td>
<td>40 19.2</td>
<td>48 39.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>41.4 13.3</td>
<td>39.3 12.7</td>
<td>37.6 11.3</td>
<td>1.5</td>
<td>n.s.</td>
<td>2.6 &lt;0.01</td>
</tr>
<tr>
<td>Age at onset of depression (years)</td>
<td>29.2 14.1</td>
<td>27.2 14.0</td>
<td>24.0 12.7</td>
<td>1.3</td>
<td>n.s.</td>
<td>3.3 &lt;0.01</td>
</tr>
<tr>
<td>Number of depressive episodes</td>
<td>23 2.0</td>
<td>2.5 2.2</td>
<td>3.1 2.7</td>
<td>0.5</td>
<td>n.s.</td>
<td>2.7 &lt;0.05</td>
</tr>
<tr>
<td>Current episode duration (weeks)</td>
<td>172.2 428.1</td>
<td>182.6 372.1</td>
<td>217.0 386.8</td>
<td>0.2</td>
<td>n.s.</td>
<td>1.0 n.s.</td>
</tr>
<tr>
<td>Global Assessment of Functioning Scale score</td>
<td>50.7 8.1</td>
<td>53.9 7.0</td>
<td>47.2 7.6</td>
<td>4.1</td>
<td>&lt;0.001</td>
<td>3.9 &lt;0.001</td>
</tr>
</tbody>
</table>

a Ratings from Schedule of Affective Disorders and Schizophrenia.
b Those not expected to work (i.e., retired, student, housewife, physically ill) were excluded from the analysis; percents are based on N=104, N=265, and N=174 for included patients, those excluded for low severity, and those excluded for suicidal ideation or comorbidity, respectively.
c Personality disorders were assessed in 391 patients; percents are based on N=60, N=208, and N=123 for included patients, those excluded for low severity, and those excluded for suicidal ideation or comorbidity, respectively.
Compared with the included patients, the patients who would be excluded because of comorbidity or suicidal ideation had more social impairment, more frequently missed work because of psychiatric reasons, were more likely to have an episode duration of greater than 2 years, experienced more previous episodes, made more suicide attempts, and were more likely to have a cluster B or cluster C personality disorder.

Discussion

When applying the exclusion criteria of recently published antidepressant efficacy trials (7) to our patients, we found that the majority would not qualify for most antidepressant efficacy trials. The 79% exclusion rate found in the present study is consistent with our previous findings. The present report extends our earlier results by demonstrating that the demographic, clinical, and psychosocial profile of patients who would be excluded from an antidepressant efficacy trial because of a comorbid anxiety or substance use disorder or substantial suicidal ideation differs from that of patients who would qualify for a trial. Specifically, we found that these excluded patients are a more chronically ill group, with more previous episodes, greater social and occupational impairment, and more personality pathology. In contrast, the psychosocial and clinical characteristics of patients who would be excluded because of insufficient symptom severity were virtually the same as the patients who would qualify for an antidepressant efficacy trial.

Antidepressant efficacy trials tend to exclude two groups of patients: those who are less likely to respond to treatment because they have substantial personality pathology and associated greater chronicity and poorer psychosocial functioning and those who are less likely to demonstrate differences in response between active medication and placebo by virtue of the milder severity of their symptoms. Although there are some studies demonstrating the efficacy of antidepressants in chronically and mildly depressed patients (8), these are relatively few. We are unaware of any large-scale placebo-controlled studies that have focused on the efficacy of treating depression in patients with personality disorders.

The results of the present study, added to the findings of studies demonstrating that a low percentage of applicants are accepted into antidepressant efficacy trials and that a low percentage of patients in clinical practice would qualify for an antidepressant efficacy trial, call for a discussion regarding how the results of antidepressant efficacy trials should be interpreted and promoted. Antidepressant efficacy trials are generally limited to patients with the greatest likelihood of demonstrating drug-placebo differences. Yet, antidepressant medications are approved and marketed without acknowledging the generalizability issue. It seems reasonable that antidepressants should be approved and marketed only for the narrow range of patients in whom they are proven effective. Approval for other groups of depressed patients should be dependent on demonstration of efficacy, analogous to requiring additional study to obtain an indication for pediatric depression or bipolar depression.

Received Feb. 17, 2004; revision received July 19, 2004; accepted Aug. 11, 2004. From the Department of Psychiatry and Human Behavior, Brown Medical School, Rhode Island Hospital, Providence. Address correspondence and reprint requests to Dr. Zimmerman, Bayside Medical Center, 235 Plain St., Providence, RI 02905; mzimmerman@lifespan.org (e-mail).

Supported in part by NIMH grants MH-48732 and MH-56404.

References

Strong Inverse Association Between Height and Suicide in a Large Cohort of Swedish Men: Evidence of Early Life Origins of Suicidal Behavior?

Patrik K.E. Magnusson, Ph.D.
David Gunnell, M.D., Ph.D.
Per Tynelius, M.S.
George Davey Smith, M.D., Ph.D.
Finn Rasmussen, M.D., Ph.D.

Objective: Previous studies have found associations between poor fetal and infant growth and the risk of suicide. The authors’ goal was to investigate the association between height—a measure of childhood growth—and suicide risk.

Method: The authors conducted a record linkage study of the birth, conscription, mortality, family, and census register data of 1,299,177 Swedish men followed from age 18 to a maximum of age 49.

Results: There were 3,075 suicides over an average follow-up period of 15 years. There was a strong inverse association between height and suicide risk. In fully adjusted models, a 5-cm increase in height was associated with a 9% decrease in suicide risk.

Conclusions: The strong inverse association between height and suicide may signify the importance of childhood exposure in the etiology of adult mental disorder or reflect stigmatization or discrimination encountered by short men in their adult lives.


H
eight is a marker of postnatal development as well as factors, including genes, that influence growth. Associations of several adult chronic diseases with height are thought to indicate the importance of the postnatal environment on disease risk (1). Few previous prospective studies have investigated the association of height with suicide. A Swedish study (2) found some evidence of a greater risk among the shortest 10% of men studied, but possible socioeconomic confounding factors were not taken into account. The other study, from South Korea (3), documented a twofold greater risk of suicide in the shortest compared with the tallest men, but this association was attenuated after the authors controlled for socioeconomic factors. A further investigation reported inverse associations between height and hospital admission for self-harm (4). In the current article, we examine the relationship between height and suicide in a large cohort of Swedish men.

Method

Swedish-born males born 1950–1981 for whom information on their biological parents was available (N=1,654,668) were identified in the Swedish Multi-Generation Register; 1,442,923 (87.2%) of these individuals had a record in the Military Service Conscription Register 1968–1999. These records were linked with the Medical Birth Register, the Cause of Death Register 1968–1999, and the Population and Housing Censuses 1970–2000.

Height measurements at age 18–19 years were obtained from the Military Service Conscription Register. The following possible confounding factors were assessed: maternal and paternal educational level (six categories); the highest socioeconomic index of either parent (four categories); body mass index; date of birth; and conscription center (six centers). Information on marital status at age 30–35 years was available for a subgroup born before 1960. For men born after 1973, information on birth weight, birth length, and gestational age were obtained from the Medical Birth Register.

Altogether, complete information was available on all variables described above for 1,299,177 (90%) of the subjects. Individuals excluded due to incomplete information were shorter (mean height=178.6 cm) than those included (mean height=179.3 cm) (t=37.05, df=1,442,921, p<0.0001) and had a higher suicide rate (age- and birth-year-adjusted hazard ratio=1.15, p<0.01).

Suicide deaths occurring between age 18 and 49 years were identified by using ICD-8 and ICD-9 codes E950–E959 or ICD-10 codes X60–X84. Associations were also investigated with 1) undetermined deaths (ICD-8 and ICD-9 codes E980–E989 or ICD-10 codes Y10–Y34) and 2) alcohol-related deaths (ICD-8 and ICD-9 codes 291, 303, 571, and E860 or ICD-10 codes F10, K70, T51, X45, X65).

We used Cox’s proportional hazards models, with age as the time axis, to investigate associations between height and suicide. Subjects were censored at their date of death, emigration, or the end of follow-up (Dec. 31, 1999). We categorized subjects into nine groups on the basis of the number of standard deviations their height was from the overall mean (≤0.0, –0.5 to ≤0.5, 0.5 to ≤1.0, 1.0 to ≤1.5, 1.5 to ≤2.0, and >2 standard deviations). We also estimated hazard ratios for 5-cm increments in height.

Results

The mean height of the 1,299,177 conscripts was 1.79 m; 711,949 (54.8%) came from white-collar families; and 3,075 (0.24%) died by suicide. Taller men had a much lower risk of suicide than shorter men (Table 1). The association was linear. A 5-cm increase in height was associated with a 9% (95% confidence interval [CI]=7%–12%) decrease in suicide risk. The effect of height changed little after adjustment for parental socioeconomic index or the participant’s body mass index.

http://ajp.psychiatryonline.org 1373
To assess whether secular increases in height may account for the observed associations, we repeated the analysis using year of birth as the time axis. The associations were unchanged: the hazard ratio per 5-cm change in height was 0.90 (95% CI=0.88–0.93) (Table 1).

Compared with height-suicide associations, height was more strongly related to undetermined deaths (hazard ratio per 5-cm increase in height=0.86, 95% CI=0.81–0.90), and alcohol-related deaths (hazard ratio=0.77, 95% CI=0.71–0.83).

In the subgroup born 1950–1965 and alive in 1990, who had had the opportunity to complete their education, the height-suicide association was only weakly attenuated—from 0.91 (95% CI=0.86–0.95) to 0.92 (95% CI=0.88–0.97)—after we controlled for education. Information on marital status at age 30–35 years was available for men born before 1960, and the fully adjusted hazard ratio in this subgroup—0.93 (95% CI=0.88–0.98)—was only weakly attenuated after adjustment for marital status—0.94 (95% CI=0.89–0.99). In the subgroup of men with birth weight data, the hazard ratio for suicide was unchanged after we controlled for birth weight, gestational age, and birth length (data not shown).

Exclusion of all conscripts with a psychiatric diagnosis at conscription (ICD-8 and ICD-9 codes 290–319) from the fully adjusted analysis did not affect the hazard ratio markedly—0.91 (95% CI=0.89–0.94). The height-suicide association was 0.94 (95% CI=0.87–1.00) among those excluded for this reason.

### Discussion

We found a twofold higher risk of suicide in short men than tall men. The associations do not appear to be attributable to socioeconomic confounding or prenatal influences on growth (5). Stronger associations were seen with alcohol-related mortality, suggesting that substance misuse may contribute to the observed patterns.

A strength of our study is that it includes 79% of all men born in Sweden in 1950–1981. It has two main limitations. We were unable to fully assess the possible influence of unemployment, relationship breakdown, or mental illness on the height-suicide associations. Our findings cannot be generalized to women or older men. Nevertheless, in Sweden, the United States, and the United Kingdom, suicides in 18–49-year-old men account for almost half of all suicides.

There are several explanations for our findings. First, psychological stress and disrupted family life in childhood impair growth (6) and may increase susceptibility to mental illness and suicidal behavior in later life (7). Short stature may be associated with a greater risk of psychosis (8), which in turn influences suicide risk. However, associations were not attenuated by excluding subjects with psychiatric diagnoses at conscription. Short individuals are more likely to be in a low social class as adults, independent of their childhood social class (9). Low social class is associated with a greater risk of suicide (10). In a subset of subjects, however, we found that educational level, a marker of socioeconomic position, had little effect on the associations. Marriage protects against suicide (11), and short individuals may be less likely to marry than taller ones (12). Marital status only weakly confounded the associations. Low weight gain in infancy may also be a risk factor for suicide in adult life (13). Finally, short children tend to have lower levels of intelligence and may suffer stigmatization and discrimination (9, 14).

Received Feb. 6, 2004; revisions received April 23 and July 15, 2004; accepted Aug. 11, 2004. From the Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Sweden; the Department of Social Medicine, University of Bristol, U.K.; and the Child and Adolescent Public Health Epidemiology Group, Department of Public Health Sciences, Karolinska Institute. Address correspondence and reprint requests to Associate Professor Rasmussen, Department of Genetics and Pathology, Uppsala University, Sweden.
Is Violent Method of Suicide a Behavioral Marker of Lifetime Aggression?

Alexandre Dumais, B.Sc.
Alain D. Lesage, M.D., M.Phil.
Aleksandra Lalovic, M.Sc.
Monique Séguin, Ph.D.
Michel Tousignant, Ph.D.
Nadia Chawky, M.Ps.
Gustavo Turecki, M.D., Ph.D.

Objective: The main purpose of this study was to investigate whether the method of suicide is a valid behavioral marker of a lifetime history of aggression.

Method: The authors applied the psychological autopsy method to investigate 310 individuals who committed suicide.

They used structured clinical assessments and personality trait scales in interviews with family members of the deceased.

Results: Violent method was associated with a higher level of lifetime aggression and a higher level of impulsivity. In addition, violent method was associated with lifetime substance abuse or dependence and psychotic disorders. Controlling for age, sex, substance disorders, and other major psychopathology, the authors found that lifetime aggression and the interaction between impulsivity and aggressive behavior remained associated with violent method.

Conclusions: These results support the use of violent method of suicide as a behavioral marker of a higher level of lifetime impulsive-aggressive behaviors.

(Am J Psychiatry 2005; 162:1375–1378)
Is Violent Method of Suicide a Behavioral Marker of Lifetime Aggression?

Alexandre Dumais, B.Sc.
Alain D. Lesage, M.D., M.Phil.
Aleksandra Lalovic, M.Sc.
Monique Séguin, Ph.D.
Michel Tousignant, Ph.D.
Nadia Chawky, M.Ps.
Gustavo Turecki, M.D., Ph.D.

Objective: The main purpose of this study was to investigate whether the method of suicide is a valid behavioral marker of a lifetime history of aggression.

Method: The authors applied the psychological autopsy method to investigate 310 individuals who committed suicide. They used structured clinical assessments and personality trait scales in interviews with family members of the deceased.

Results: Violent method was associated with a higher level of lifetime aggression and a higher level of impulsivity. In addition, violent method was associated with lifetime substance abuse or dependence and psychotic disorders. Controlling for age, sex, substance disorders, and other major psychopathology, the authors found that lifetime aggression and the interaction between impulsivity and aggressive behavior remained associated with violent method.

Conclusions: These results support the use of violent method of suicide as a behavioral marker of a higher level of lifetime impulsive-aggressive behaviors.

Suicide is an important public health problem (1). The method of suicide does not appear to be randomly distributed. For example, violent method of suicide is more often used by males than females (2–4), and violent method is more common in suicide completers affected by psychosis (5). On the other hand, the relationship between violent method and age remains unclear (4–7). Since the early work of Åsberg (8), suicide method has been used in neubiological studies of suicide attempters to select subjects who are more likely to have low indexes of serotonergic neurotransmission, which in turn tend to correlate with higher levels of aggression (9). However, the relation-
TABLE 1. Axis I and Axis II Disorders and Impulsive and Aggressive Behaviors in Individuals Who Committed Suicide by Violent or Nonviolent Methods

<table>
<thead>
<tr>
<th>Disorder or Behavior</th>
<th>Nonviolent Method of Suicide</th>
<th>Violent Method of Suicide</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Axis I disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Last 6 months</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety disordera,b</td>
<td>156</td>
<td>4</td>
<td>12.9</td>
</tr>
<tr>
<td>Mood disorderc</td>
<td>156</td>
<td>21</td>
<td>67.7</td>
</tr>
<tr>
<td>Alcohol and/or drug abuse or dependence</td>
<td>156</td>
<td>64</td>
<td>29.0</td>
</tr>
<tr>
<td>Psychotic disorderc</td>
<td>156</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Lifetime</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety disordera,b</td>
<td>156</td>
<td>4</td>
<td>12.9</td>
</tr>
<tr>
<td>Mood disorderc</td>
<td>156</td>
<td>14</td>
<td>45.2</td>
</tr>
<tr>
<td>Alcohol and/or drug abuse or dependence</td>
<td>156</td>
<td>8</td>
<td>25.8</td>
</tr>
<tr>
<td>Psychotic disorderc</td>
<td>156</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Axis II disordersd</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster A: schizoid, paranoid</td>
<td>157</td>
<td>1</td>
<td>3.1</td>
</tr>
<tr>
<td>Cluster B: borderline, antisocial, histrionic, narcissistic</td>
<td>157</td>
<td>10</td>
<td>31.3</td>
</tr>
<tr>
<td>Cluster C: dependent, obsessive-compulsive, avoidant, depressive, passive-aggressive</td>
<td>157</td>
<td>5</td>
<td>16.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Impulsive and aggressive behaviors</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barratt Impulsivity Scale total score</td>
<td>64.01</td>
<td>15.22</td>
<td>69.33</td>
<td>14.49</td>
<td>-1.89</td>
<td>144</td>
<td>0.06</td>
</tr>
<tr>
<td>Brown-Goodwin Lifetime History of Aggression total score</td>
<td>6.44</td>
<td>5.80</td>
<td>13.10</td>
<td>13.87</td>
<td>-2.97</td>
<td>106</td>
<td>0.004</td>
</tr>
</tbody>
</table>

a Numbers of subjects in analysis vary according to the particular analysis listed in each row because information was not available for all subjects on all measures. The degree of overlap between measures is variable. Percents are based on N=31 or N=32 for subjects who committed suicide by violent methods and N=125 for subjects who committed suicide by nonviolent methods.

b Panic disorder, agoraphobia without panic disorder, social phobia, specific phobia, obsessive-compulsive disorder, generalized anxiety disorder, anxiety disorder not otherwise specified.

c Fisher's exact test.

d Major depression, depression not otherwise specified.

e Schizophrenia, schizoaffective disorder, delusional disorder.

Method

Subjects for this study were individuals who had committed suicide identified through the Coroner's Office and collected sequentially from the Greater Montreal area. Seventy-three percent of the families of these individuals agreed to participate in our clinical study. After a period averaging 4 months, these families were contacted again for interviews. This study was approved by our local institutional review board, and written informed consent was obtained from all participating families.

Suicides were classified on the basis of information provided by the Coroner according to criteria used in previous studies (2, 5, 12). Overdoses, poisoning, gas, and drowning were classified as nonviolent methods of suicide; all other methods were classified as violent.

Psychiatric diagnoses in individuals who committed suicide were made by means of the psychological autopsy method. This technique has been well validated and is outlined elsewhere (13–15). Psychiatric diagnoses were obtained by using the Structured Clinical Interview (SCID) (16) for DSM-IV axis I diagnoses and axis II personality disorders. Before applying the SCID, we used the Schedule for Affective Disorders and Schizophrenia for School-Age Children (17) modified to include questions assessing personality disorders adapted from the Children's Depression Inventory (18). As reported elsewhere (19), diagnoses obtained using these two different methods had an excellent concordance rate. Information collected with these interviews, from the Coroner's notes, and from medical records was used by a panel of clinicians who made consensus DSM-IV diagnoses.

Two or more interviewers were asked to rate the same subject separately, and kappa coefficients for key diagnoses were excellent: kappa=0.96 for major depression, kappa=0.98 for alcohol abuse/dependence, kappa=1.0 for drug abuse/dependence, kappa=1.0 for bipolar disorder, kappa=1.0 for schizophrenia, and kappa=1.0 for cluster B personality disorders. These data are consistent with those previously presented by our group (20) and probably are a direct result of frequent training sessions to avoid drifting between interviewers.

The Brown-Goodwin Lifetime History of Aggression (21) and the Barratt Impulsivity Scale (22) were used to assess lifetime aggressive and impulsive behaviors, respectively. The Temperament and Character Inventory (23) was used to complement this information. Internal consistency estimates were excellent overall: alpha=0.88 for the informant version of the Brown-Goodwin Lifetime History of Aggression, alpha=0.89 for the Barratt Impulsivity Scale, and alpha between 0.73 and 0.88 for the Temperament and Character Inventory.

In addition, we compared information obtained using two different informants for the same deceased individual for personal-
ity trait measures and observed no significant differences (p values between 0.25 and 0.94). Furthermore, in studies with living subjects we compared information obtained with an informant and the subject and found no significant differences (p values between 0.67 and 0.98).

Chi-square, odds ratios, and Fisher's exact tests were used to compare categorical variables, and t tests were used in the analysis of continuous variables. Logistic regression was used to obtain adjusted risks.

Results

A total of 310 individuals who had committed suicide were investigated: 36 female subjects and 274 male subjects with a mean age of 39.45 years (SD=13.91). Two hundred forty-two (78.1%) of these individuals used a violent method: 166 (68.6%) hanged themselves, 39 (16.1%) used a firearm, 12 (5.0%) used laceration, 18 (7.4%) jumped from a height, and seven (2.9%) used one of the following: traffic accident, electrocution, self-immolation, or strangulation with a plastic bag. On the other hand, 68 (21.9%) of the individuals who committed suicide used nonviolent methods. Of these, 36 (52.9%) died intoxicated by gases, 22 (32.4%) by drug poisoning, and 10 (14.7%) by drowning. This distribution of suicide methods is consistent with distribution of suicide methods is consistent with the Quebec Coroner's Office in 2000. The average age was lower in the violent method group (38.20 compared with 43.96 years) (t=2.54, df=305, p<0.02), and female subjects showed a tendency to use a nonviolent method (χ²=3.09, df=1, p<0.08).

Last-6-month prevalence rates of anxiety disorders, mood disorders, and alcohol and/or drug problems (abuse/dependence) were comparable between groups. Similarly, lifetime prevalence rates of anxiety disorders and mood disorders were comparable between groups, as were the prevalence rates of cluster A, cluster B, and cluster C personality disorders. In contrast, prevalence rates of lifetime alcohol and/or drug problems and psychotic disorders (6-month and lifetime) were significantly associated with violent method of suicide (Table 1).

Measures of lifetime history of aggressive behaviors were higher in the group that used a violent method of suicide. In addition, we found a nonsignificantly higher level of impulsive behaviors as measured by the Barratt Impulsivity Scale in the violent method group. Aggressive and impulsive behavior scores were also found to be significantly correlated (r=0.45, N=124, p<0.001). On the other hand, Temperament and Character Inventory measures were not statistically different between groups (data not shown). Finally, when we controlled for the effect of age, sex, substance disorders, and psychopathology (major psychiatric axis I disorders as a categorical variable), we found that history of lifetime aggression (p=0.03) and the interaction between lifetime aggression and lifetime impulsivity (p<0.06) remained associated with a violent method of suicide.

Discussion

To our knowledge, this study is the first to investigate the relationship between measures of aggression, impulsivity, and suicide method in individuals who committed suicide. Overall, our analysis supports the use of suicide method as a possible behavioral marker of lifetime impulsive aggression.

We found that younger age was associated with violent method and observed a nonsignificant gender effect. The effect of gender is consistent with previous findings (2, 4, 6), and its nonsignificance may be a consequence of the relatively small number of female subjects included in this study. Although suicide cases in this study were obtained sequentially, our group has been focusing on male suicide and only more recently has started recruiting females. The age effects found in this study support the finding by Cowell et al. (7) that older individuals who commit suicide are more likely to use nonviolent methods. Finally, we observed that lifetime alcohol and/or drug problems and psychotic disorders were associated with use of violent methods. In addition, significant correlations between lifetime history of substance misuse and impulsive behavior (r=0.41, N=135, p=0.001) and aggression (r=0.32, N=124, p=0.004) were found. These findings are consistent with those of previous studies (24, 25) suggesting that there is probably a relationship between chronic alcohol and/or drug consumption and lifetime aggression and violence. Nevertheless, when we controlled for the history of substance use, as well as for other positive variables, we found that higher levels of impulsive and aggressive behavior and their interaction remained significant predictors of violent methods of suicide.

Given that the lethality associated with violent methods is considerably higher than that associated with nonviolent methods (26), a possible hypothesis that could be drawn from our results is that the observed excess of impulsive-aggressive behavior among individuals who committed suicide may be a direct consequence of the fact that these individuals were more likely to use a violent method. This hypothesis is supported by recent data suggesting that differences in method explain a substantial amount of the difference between completed suicide and medically serious nonfatal suicide attempts (27).

The major limitation of this study is intrinsic to the psychological autopsy method. However, our results on validity are consistent with the literature (13–15) and support the use of behavioral assessments obtained by means of informants. Other limitations in our analyses include the infrequency of diagnoses such as psychotic disorders.

In conclusion, our results suggest that violent method of suicide could be a valid behavioral marker of lifetime impulsive aggression. This relationship is probably mediated by several other factors that this study did not aim to assess and that remain to be investigated further. Additional research including suicide attempters with a larger num-

http://ajp.psychiatryonline.org

Am J Psychiatry 162:7, July 2005
ber of female subjects who committed suicide and more individuals who committed suicide by nonviolent methods is needed to confirm and further understand the relationship between impulsive-aggressive behaviors, suicide method, and suicide outcome.

References

1. Suicide in the United States. Atlanta, Centers for Disease Control and Prevention, 2003
11. Muller-Oerlinghausen B, Roggenbach J: Concretion in biological suicide research—are we eating the menu instead of the meal? some thoughts on present research strategies. Pharmacopsychiatry 2002; 35:44–49
Prediction of Panic Response to a Respiratory Stimulant by Reduced Orbitofrontal Cerebral Blood Flow in Panic Disorder

Justine M. Kent, M.D.
Jeremy D. Coplan, M.D.
Osama Mawlawi, Ph.D.
Jose M. Martinez, M.A.
Susan T. Browne, B.A.
Mark Slifstein, Ph.D.
Diana Martinez, M.D.
Anissa Abi-Dargham, M.D.
Marc Laruelle, M.D.
Jack M. Gorman, M.D.

Objective: Lack of appropriate top-down governance by frontal cortical regions over a hypersensitive amygdala-centered fear neurocircuitry has been hypothesized to be central in the pathophysiology of panic disorder. The aim of this study was to examine regional cerebral blood flow changes in response to anxiety/panic provocation in subjects with panic disorder and healthy comparison subjects.

Method: Quantitative water method positron emission tomography was used to obtain brain images of five untreated subjects with panic disorder and five healthy comparison subjects before and during anxiogenic challenge with intravenous doxapram, an acute respiratory stimulant.

Results: Baseline perfusion of the orbitofrontal cortex predicted panic attacks: lower perfusion was associated with heightened anxiety in response to doxapram challenge.

Conclusions: The orbitofrontal cortex may be important in the regulation of responding to fear and is a potential area of aberrant functioning in panic disorder.

Although much is now known about the pathophysiology of panic disorder, the neurocircuitry central to panic is still being delineated. The role of inhibitory cortical inputs on amygdalofugal pathways in the development of panic anxiety is especially poorly understood.

Although the results are not completely consistent, the majority of functional imaging studies in panic disorder support frontal cortical deactivation during panic anxiety provoked by pharmacological challenges (1–3) and during spontaneous panic attacks (4). Because the medial and orbitofrontal regions of the prefrontal cortex are known to have extensive connections with the amygdala, exerting a primarily inhibitory influence on its activity (5), they are of prime interest as areas of potentially aberrant functioning in panic disorder.

The aims of this study were to use quantitative water method ([15O]H2O) positron emission tomography (PET) to test the following hypotheses: 1) Subjects with panic disorder demonstrate greater activation of the amygdala than healthy comparison subjects in response to an anxiety-provoking challenge. 2) Subjects with panic disorder demonstrate less cortical restraint than healthy comparison subjects over subcortical fear pathway structures, as manifest by hypoactivity in medial and orbitofrontal cortical regions.

Method

Five subjects with a DSM-IV diagnosis of panic disorder (four female, one male; mean age=27.2 years, SD=3.2) and five healthy comparison subjects (three female, two male; mean age=31.6 years, SD=6.8) participated in the study. All subjects participated in a psychiatric interview and the Structured Clinical Interview for DSM-IV Axis I Disorders (6), and all were determined to be in good physical health after a complete medical evaluation. Subjects with panic disorder were included if they met DSM-IV criteria for panic disorder; no other current axis I diagnoses were allowed except dysthymia. Subjects with panic disorder were excluded if they had a lifetime history of schizophrenia or bipolar disorder. Healthy comparison subjects did not meet criteria for any current or lifetime history of an axis I disorder. All subjects were required to be free of psychotropic medications for 6 weeks before scanning. The study was approved by the institutional review boards of New York State Psychiatric Institute and Columbia University. After complete description of the study to the subjects, written informed consent was obtained from all who participated.

Subjects were scanned with [15O]H2O PET on an ECAT EXACT HR+ PET scanner (Siemens CTI, Knoxville, Tenn.); input function was measured with arterial blood samples. The respiratory stimulant doxapram was used to induce anxiety in subjects (7). The pharmacological action of doxapram hydrochloride is mediated through the peripheral carotid chemoreceptors, producing respiratory stimulation within 20–40 seconds of injection, with a peak effect at 1–2 minutes and a total duration of action of approximately 5 minutes. Because the effects of doxapram preclude a double-blind randomized study design, a single-blind nonrandomized design was used.

All subjects underwent five activations (scans), each consisting of an intravenous bolus injection of 20 mCi [15O]H2O. Scans were acquired 15 minutes apart in a fixed order: resting baseline, repeat resting baseline, placebo injection, placebo injection, doxapram injection (0.5 mg/kg). Scanning began 30 seconds after placebo and doxapram injections. Throughout the scanning session, repeated anxiety assessments were made with three scales: 1) Acute Panic Inventory (8), 2) 10-point Anxiety Scale (8), and 3) 10-point Borg Breathlessness Scale (9). Occurrence of panic was judged on the basis of DSM-IV criteria of a crescendo of fear/anxiety with four or more associated symptoms (rated on the Acute Panic Inventory).

http://ajp.psychiatryonline.org

Am J Psychiatry 162:7, July 2005
FIGURE 1. Relationship of Baseline Orbitofrontal Cerebral Blood Flow (CBF) to Scores on the Acute Panic Inventory During Doxapram Challenge in Four Patients With Panic Disorder and Five Healthy Comparison Subjects

Dynamic PET imaging data were summed and coregistered with magnetic resonance imaging (MRI). Data were fitted to the integrated form of the Kety-Schmidt equation (10) according to a lookup table. Summed PET data (from 30 seconds to 2 minutes after injection) were matched to flow values from a lookup table generated by fitting the equations to a range of possible summed activity values. The delay between the time of measurement of arterial tracer concentration and the time of brain exposure was incorporated by fitting whole brain data to the dynamic form of the equations and including delay time as a fitted parameter. Parametric maps were generated, with one flow value for each voxel. Regions of interest were drawn on individual patients’ MRIs and transferred to the blood flow maps, generating regional mean flow values. Subcortical regions of interest included the midbrain, hippocampus, amygdala, thalamus, and striatum. The prefrontal cortex was divided into the following regions of interest: anterior cingulate, medial, dorsolateral, orbitofrontal, and subgenual.

Change in cerebral blood flow (CBF) over time (scans 1–5) was analyzed by using repeated-measures analysis of variance. Behavioral data were compared with regional CBF (rCBF) values from the scans by using correlational analyses based on the a priori hypotheses described above.

Results

All five of the subjects with panic disorder panicked in response to doxapram, compared with one of the five healthy subjects, a significant difference ($\chi^2$=5.76, df=1, $p<0.02$). One of the subjects with panic disorder was dropped from the PET data analysis because of a technical problem with processing her PET data (this subject was 22 years old). There were no significant differences in rCBF in the amygdala in response to placebo or doxapram injections between the subjects with panic disorder and the comparison subjects. Similarly, there were no significant differences between the groups in global (weighted average) subcortical or cortical blood flow in response to placebo or doxapram injections and no significant differences between the groups when regions of interest were analyzed individually (Bonferroni correction for multiple comparisons).

When all subjects were considered together (N=9), baseline orbitofrontal CBF distinguished subjects who panicked (N=5) from those who did not panic (N=4) in response to doxapram injection ($t=2.82$, df=7, $p<0.05$). Additionally, baseline orbitofrontal CBF was negatively correlated with anxiety scores on the Acute Panic Inventory (r=−0.83, N=9, $p<0.05$) (Figure 1), the 10-point Anxiety Scale (r=−0.77, N=9, $p<0.05$), and the Borg Breathlessness Scale (r=−0.75, N=9, $p<0.05$) in response to doxapram injection.

Conclusions

Perfusion of the orbitofrontal prefrontal cortex at baseline predicted panic vulnerability to respiratory challenge across subjects, suggesting that input from this cortical region may be important in suppressing fear responding. In addition, orbitofrontal blood flow at baseline correlated negatively with scores on the Acute Panic Inventory, 10-point Anxiety Scale, and Borg Breathlessness Scale in response to doxapram challenge: higher levels of CBF in the orbitofrontal region was associated with lower anxiety scores and less breathlessness.

These data should be viewed cautiously given the small number of subjects and the high risk of a type II error. This may account for our inability to demonstrate a relationship between panic anxiety and amygdala activation. Lack of differentiation between the panic and comparison groups in rCBF response to the anxiogenic challenge was perhaps due to a “floor” effect, in which the respiratory response to doxapram, resulting in hyperventilation, hypocapnia, and resultant physiological vasoconstriction, overshadowed potential differences between groups. This potential confound of hypocapnia-induced vasoconstriction could be controlled for by measuring PCO2 with serial blood gases in future studies. Hypocapnia-induced vasoconstriction, however, is likely to be a global effect that would not account for the regional difference identified in the orbitofrontal cortex at baseline. Despite the limitations of this study, these results may stimulate more studies focusing on the orbitofrontal cortex and directed at determining the nature of inhibitory cortical input to the amygdala in the development of panic vulnerability.

Received Oct. 23, 2003; revision received June 15, 2004; accepted July 7, 2004. From the Department of Psychiatry and Department of Radiology, Columbia University College of Physicians & Surgeons; New York State Psychiatric Institute, New York; State University of New York-Downstate, Brooklyn, N.Y.; and Mount Sinai School of Medicine, New York. Address correspondence and reprint requests to Dr. Kent, Department of Psychiatry, Unit 41, Columbia University, 1051 Riverside Dr., New York, NY 10032; jmk14@columbia.edu (e-mail). Supported by Eli Lilly & Co.
Acute Stress Disorder and Posttraumatic Stress Disorder in Children and Adolescents Involved in Assaults or Motor Vehicle Accidents

Richard Meiser-Stedman, Ph.D.
William Yule, Ph.D.
Patrick Smith, Ph.D.
Ed Glucksman, F.F.A.E.M.
Tim Dalgleish, Ph.D.

Objective: The authors investigated acute stress disorder and later posttraumatic stress disorder (PTSD) in children and adolescents who had been involved in assaults or motor vehicle accidents.

Method: They interviewed 93 patients 10–16 years old who were seen in an emergency department for having been assaulted or involved in a motor vehicle accident within 4 weeks after the assault or accident to assess acute stress disorder. At 6 months, they reinterviewed 64 (68.8%) of the patients to assess PTSD.

Results: At initial interview, 18 (19.4%) of the 93 patients had acute stress disorder and 23 (24.7%) met all acute stress disorder criteria except dissociation. At 6 months, eight of the 64 patients (12.5%) had PTSD. Acute stress disorder and PTSD did not differ in prevalence between patients who had been assaulted and those who had been in accidents. Sensitivity and specificity statistics and regression modeling revealed that the diagnosis of acute stress disorder was a good predictor of later PTSD but that dissociation did not play a significant role.

Conclusions: Acute stress disorder has merit as a predictor of later PTSD in children and adolescents, but dissociation has questionable utility.

References

Psychological trauma occurs at high rates in children and adolescents (1), and posttraumatic stress disorder (PTSD) in this age group has attracted considerable clinical and research interest. However, the diagnosis of acute stress disorder, introduced in DSM-IV, has received relatively little attention in younger populations (2), despite a growing body of research in adults (3). Unlike PTSD, which is diagnosed at least 4 weeks after trauma, acute stress disorder is diagnosed 2 days to 4 weeks after trauma. Acute stress disorder also differs from PTSD in being explicitly conceived as a dissociative response to trauma requiring at least three of a possible five dissociation symptoms. An important public health marker of the utility of acute stress disorder is its ability to predict later PTSD, thus allowing clinicians to focus resources on susceptible individuals (4). In adults, acute stress disorder is a good predictor of later PTSD (5, 6), but the dissociation symptoms appear to add little (7).

Several studies have examined acute stress disorder symptoms in younger populations (8–12). However, only one study (11) examined the power of the acute stress disorder diagnosis (derived solely from questionnaire re-
Acute Stress Disorder and Posttraumatic Stress Disorder in Children and Adolescents Involved in Assaults or Motor Vehicle Accidents

Richard Meiser-Stedman, Ph.D.
William Yule, Ph.D.
Patrick Smith, Ph.D.
Ed Glucksman, F.F.A.E.M.
Tim Dalgleish, Ph.D.

Objective: The authors investigated acute stress disorder and later posttraumatic stress disorder (PTSD) in children and adolescents who had been involved in assaults or motor vehicle accidents.

Method: They interviewed 93 patients 10–16 years old who were seen in an emergency department for having been assaulted or involved in a motor vehicle accident within 4 weeks after the assault or accident to assess acute stress disorder. At 6 months, they reinterviewed 64 (68.8%) of the patients to assess PTSD.

Results: At initial interview, 18 (19.4%) of the 93 patients had acute stress disorder and 23 (24.7%) met all acute stress disorder criteria except dissociation. At 6 months, eight of the 64 patients (12.5%) had PTSD. Acute stress disorder and PTSD did not differ in prevalence between patients who had been assaulted and those who had been in accidents. Sensitivity and specificity statistics and regression modeling revealed that the diagnosis of acute stress disorder was a good predictor of later PTSD but that dissociation did not play a significant role.

Conclusions: Acute stress disorder has merit as a predictor of later PTSD in children and adolescents, but dissociation has questionable utility.

(Psychological trauma occurs at high rates in children and adolescents (1), and posttraumatic stress disorder (PTSD) in this age group has attracted considerable clinical and research interest. However, the diagnosis of acute stress disorder, introduced in DSM-IV, has received relatively little attention in younger populations (2), despite a growing body of research in adults (3). Unlike PTSD, which is diagnosed at least 4 weeks after trauma, acute stress disorder is diagnosed 2 days to 4 weeks after trauma. Acute stress disorder also differs from PTSD in being explicitly conceived as a dissociative response to trauma requiring at least three of a possible five dissociation symptoms. An important public health marker of the utility of acute stress disorder is its ability to predict later PTSD, thus allowing clinicians to focus resources on susceptible individuals (4). In adults, acute stress disorder is a good predictor of later PTSD (5, 6), but the dissociation symptoms appear to add little (7).

Several studies have examined acute stress disorder symptoms in younger populations (8–12). However, only one study (11) examined the power of the acute stress disorder diagnosis (derived solely from questionnaire re-
TABLE 1. Sensitivity, Specificity, and Positive and Negative Predictive Power of Acute Stress Disorder, Subacute Stress Disorder, and Early PTSD Criteria and Diagnoses to Predict Later PTSD Among Children and Adolescents Involved in Assaults or Motor Vehicle Accidents

<table>
<thead>
<tr>
<th>Criterion/Diagnosis</th>
<th>Frequency</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive Predictive Power</th>
<th>Negative Predictive Power</th>
<th>Cases Correctly Allocated (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute stress disorder criteria</td>
<td>N=64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Traumatic stressor</td>
<td>67</td>
<td>72.0</td>
<td>1.00</td>
<td>0.32</td>
<td>0.17</td>
<td>1.00</td>
</tr>
<tr>
<td>B. Dissociation (at least three symptoms)</td>
<td>50</td>
<td>53.8</td>
<td>0.75</td>
<td>0.55</td>
<td>0.19</td>
<td>0.94</td>
</tr>
<tr>
<td>C. Reexperiencing (at least one symptom)</td>
<td>63</td>
<td>67.7</td>
<td>1.00</td>
<td>0.30</td>
<td>0.17</td>
<td>1.00</td>
</tr>
<tr>
<td>D. Avoidance (at least one symptom)</td>
<td>63</td>
<td>67.7</td>
<td>0.88</td>
<td>0.34</td>
<td>0.16</td>
<td>0.95</td>
</tr>
<tr>
<td>E. Arousal (at least one symptom)</td>
<td>75</td>
<td>80.6</td>
<td>1.00</td>
<td>0.23</td>
<td>0.16</td>
<td>1.00</td>
</tr>
<tr>
<td>F. Impairment</td>
<td>36</td>
<td>38.7</td>
<td>0.75</td>
<td>0.64</td>
<td>0.23</td>
<td>0.95</td>
</tr>
<tr>
<td>Acute stress disorder diagnosis (A + B + C + D + E + F)</td>
<td>18</td>
<td>19.4</td>
<td>0.50</td>
<td>0.88</td>
<td>0.36</td>
<td>0.92</td>
</tr>
<tr>
<td>Subacute stress disorder diagnosis (A + C + D + E + F)</td>
<td>23</td>
<td>24.7</td>
<td>0.63</td>
<td>0.82</td>
<td>0.33</td>
<td>0.94</td>
</tr>
<tr>
<td>Early PTSD criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reexperiencing (at least one symptom)</td>
<td>69</td>
<td>74.2</td>
<td>1.00</td>
<td>0.25</td>
<td>0.16</td>
<td>1.00</td>
</tr>
<tr>
<td>Avoidance (at least three symptoms)</td>
<td>49</td>
<td>52.7</td>
<td>0.50</td>
<td>0.45</td>
<td>0.11</td>
<td>0.86</td>
</tr>
<tr>
<td>Arousal (at least two symptoms)</td>
<td>60</td>
<td>64.5</td>
<td>0.88</td>
<td>0.38</td>
<td>0.17</td>
<td>0.95</td>
</tr>
<tr>
<td>Early PTSD diagnosis (early PTSD criteria + A + F)</td>
<td>23</td>
<td>24.7</td>
<td>0.63</td>
<td>0.79</td>
<td>0.29</td>
<td>0.94</td>
</tr>
<tr>
<td>PTSD diagnosis at 6-month assessment (N=64)</td>
<td>8</td>
<td>12.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Sensitivity=the probability that someone diagnosed with PTSD at follow-up had a given diagnosis, etc., at initial interview; Specificity=the probability that someone not diagnosed with PTSD at follow-up did not have a given diagnosis, etc., at initial interview; positive predictive power=the probability that someone who has a given diagnosis, etc., goes on to have a diagnosis of PTSD at 6-month follow-up; and negative predictive power=the probability that someone who does not have a given diagnosis, etc., does not go on to have a diagnosis of PTSD at 6-month follow-up.

Sponsoring this study were the National Institute of Mental Health (Grant MH-60850) and the Youth and Family Violence Research Institute, London, Ont., Canada. The authors wish to thank彩色图像中的参与者提供了他们的同意和参与。To our knowledge, the present study is the first to assess acute stress disorder in children and adolescents using structured clinical interview. Participants were assessed 2–4 weeks and 6 months after physical assault or motor vehicle accident. These events were selected because both are common, single-incident traumas with comparable rates of acute stress disorder in adults (5, 6). Our goal was to compare the utility of diagnoses of acute stress disorder, subacute stress disorder (acute stress disorder minus the dissociation criterion) (3), and “early PTSD” (PTSD without the duration criterion) (7) in predicting later PTSD.

Method

Children and adolescents (10–16 years old) who were treated in a London emergency department following motor vehicle accident or assault met study criteria (N=343). One hundred nineteen (34.7%) of these patients could not be contacted because of incomplete emergency department records, two (0.6%) were immediately referred for treatment, and 116 (33.8%) declined to participate. The 106 children and adolescents (30.9%) who consented to participate were assessed within 4 weeks. Of these, 93 (87.7%) completed an initial clinical interview and 64 (60.4%) completed a second interview at 6 months. The mean age of the 93 patients who completed the initial interview was 13.9 years (SD=1.9); 33 were girls.

There was no difference between participants and nonparticipants in terms of sex, type of trauma, injury severity, or ethnicity, although participants were significantly younger (nonparticipants’ mean age=14.8, SD=1.9 t=1.98, df=364.1, p<0.05). Participants reinterviewed at 6 months were no more or less likely than participants not reinterviewed to meet criteria for initial acute stress disorder. After description of the study, written informed consent from both children and their caregivers was obtained for all participants.

The child version of the Anxiety Disorders Interview Schedule for DSM-IV (13) was used to diagnose acute stress disorder and PTSD. This instrument is a structured interview schedule with good psychometric properties (13, 14) for the assessment of anxiety disorders in young subjects. It does not include the dissociative symptoms of acute stress disorder. Therefore, we designed a number of supplementary interview items (available from R.M.-S.) to assess these symptoms, guided by DSM-IV and existing adult acute stress disorder interview schedules (15).

Subacute stress disorder was defined as the full acute stress disorder diagnosis minus the dissociation criterion. Early PTSD was defined as PTSD at initial assessment minus the duration criterion (7). Internal reliabilities for acute stress disorder, subacute stress disorder, and early PTSD diagnoses in the current study were high (Cronbach’s alpha=0.85–0.87). There was unanimous diagnostic agreement between independent raters for 11 initial and 10 follow-up interviews (kappa=1.00).

Results

Table 1 shows how many children and adolescents met criteria for acute stress disorder, subacute stress disorder, and early and later PTSD. Chi-square analyses revealed no significant differences in prevalence for any diagnosis between those exposed to assaults and those involved in motor vehicle accidents.

Table 1 also shows the sensitivity, specificity, and positive and negative predictive power of the individual symptom criteria and the acute stress disorder, subacute stress disorder, and early PTSD diagnoses at initial interview to predict later PTSD, as well as the number of later PTSD cases that were correctly diagnosed by each criterion or
diagnosis. Subacute stress disorder was the diagnosis that gave the best balance of sensitivity and specificity.

We used logistic regression to examine whether the acute stress disorder dissociation criterion explained any unique variance in later PTSD. Subacute stress disorder was entered in the first step, resulting in a significant model ($\chi^2=6.56$, df=1, p<0.01) and accounting for unique variance ($\chi^2=6.33$, df=1, p<0.01). Entering the dissociation criterion in the second step did not significantly improve the model's ability to predict later PTSD and did not account for any unique variance.

Discussion

The rate of acute stress disorder in this group of children and adolescents (19.4%) is similar to that found in adults (5, 6) but slightly higher than in other child studies (8, 11). As in adult studies, acute stress disorder occurred at similar rates among subjects involved in assaults or motor vehicle accidents (5, 6).

The acute stress disorder diagnosis was a good predictor of later PTSD at follow-up, correctly classifying 82.8% of PTSD cases. However, subacute stress disorder provided a better balance between sensitivity and specificity at predicting later PTSD than full acute stress disorder. Furthermore, regression analysis revealed that the dissociation criterion did not significantly enhance the ability of subacute stress disorder to predict later PTSD. As in the adult literature (7), acute stress disorder and early PTSD were equally effective predictors of later PTSD.

The study had two important limitations. First, the relatively small number of subjects may have weakened statistical power. Second, the study group comprised older children (essentially preadolescents) and adolescents. Studies examining acute stress disorder in larger and younger populations are therefore needed.

In conclusion, the acute stress disorder diagnosis is a good predictor of later PTSD in children and adolescents. However, our data indicate no unique role for the dissociation criterion of acute stress disorder in these patients.

The authors thank Nicola Leete and other staff at King's College Hospital, London, for support in recruiting participants and Dr. Sean Perrin for training given to Dr. Meiser-Stedman.

References

48,XXYY Syndrome, Mood Disorder, and Aggression

To the Editor: A man with 48,XXYY syndrome was evaluated for suicidal and aggressive behavior. This unusual syndrome alerted our service to the associated psychopathology in this genetic disorder.

Mr. A, a 24-year-old man with a documented chromosomal abnormality of 48,XXYY evaluated by a standard cytogenetics technique, was admitted to the hospital for suicidal ideation and aggressive behavior toward his brother, whom he pushed down the stairs. He endorsed symptoms suggestive of a mixed bipolar episode before admission. He appeared to have borderline intelligence, with a documented IQ of 70–80, and was able to finish high school. There was no family history of psychiatric or genetic disorders. His physical appearance was notable for a tall stature, gynecomastia, and truncal obesity. His mental status examination was significant for irritability and a labile, intense affect. His cognition was intact. He had minimal insight into his illness.

Upon admission, Mr. A was given olanzapine and was later switched to oxcarbazepine and citalopram. Provisional diagnoses of bipolar disorder, not otherwise specified, and cluster B personality traits were made. Mr. A was initially uncooperative and frequently became agitated. Toward the end of admission, Mr. A seemed to respond to oxcarbazepine. His mood and behavior improved. He was discharged after 10 days of hospitalization with oxcarbazepine, 1200 mg/day, and citalopram, 40 mg/day.

48,XXYY syndrome was initially considered a variant of Klinefelter’s syndrome (1). Nowadays, it is accepted as a distinct clinical and genetic entity (2, 3). Individuals with this syndrome are more aggressive, more intellectually handicapped, and taller than people with Klinefelter’s syndrome (4). Children and adolescents with 48,XXYY syndrome often come to psychiatrists for behavioral problems, sometimes even before the chromosomal diagnosis is made. The incidences among newborn and institutionalized mentally retarded patients are 1/50,000 and 0.08%–0.33%, respectively (5). Clinical signs of 48,XXYY are generally nonspecific. Some of the common features include tall stature, gynecomastia, truncal obesity, skin ulcers, and a craniofacial dysmorphism described as a ‘pugilistic’ facial appearance (7).

To our knowledge, there are no reports of using psychotropic drugs specifically to treat behavioral problems in this patient population.

References

Dysbetalipoproteinemia and Clomipramine

To the Editor: Approximately 1% of North Americans and Northern Europeans are homozygous for apolipoprotein-E2 (APOE-2) (1). However, overt hyperlipoproteinemia occurs in only 2%–10% of these people, resulting in a markedly increased risk of atherosclerosis (1). There are numerous environmental factors known to precipitate hyperlipoproteinemia in APOE-2 homozygotes. These include the coexistence of other genetic dyslipidemias, hypothyroidism, diabetes mellitus, menopause, alcohol, poor diet, and obesity (1). Although protease inhibitors have been observed to precipitate dysbetalipoproteinemia (2), there are no such reports implicating tricyclic antidepressants. However, mild (3, 4) to moderate (5) increases induced by tricyclic antidepressants in total cholesterol have been observed in the general population with these medications.

Mr. A, a 48-year-old man with major depressive disorder, was referred to our clinic for severe dyslipidemia. At the time, he was not taking lipid-lowering medications. For most of the previous 20 years, he had been taking the tricyclic antidepressant clomipramine, first at 75 mg/day and then at 150 mg/day. Attempts to control his symptoms with selective serotonin reuptake inhibitors had failed. His other medications were carbamazepine, 200 mg t.i.d., and clonazepam, 1.5 mg at bedtime.

Mr. A’s family physician had described his initial lipid elevations as “moderate” and discovered them about 5 years after the initiation of clomipramine. However, the increase to 150 mg/day resulted in severe dyslipidemia: a total cholesterol level of 694.9 mg/dl, a triglyceride level of 637.2 mg/dl, a high-density lipoprotein cholesterol level of 81.1 mg/dl, and a total cholesterol/high-density lipoprotein cholesterol ratio of 8.6. Of interest, before his referral, Mr. A had implicated clomipramine by discontinuing it under his psychiatrist’s supervision. Dramatic improvements were observed after 11 days. His total cholesterol level was 338.2 mg/dl, his triglyceride level was 210.6 mg/dl, his high-density lipoprotein cholesterol level was 51.0 mg/dl, his low-density lipoprotein cholesterol level was 245.6 mg/dl, and his total cholesterol/high-density lipoprotein cholesterol ratio was 6.6.

Within weeks, Mr. A’s depressive symptoms returned, and clomipramine was reinitiated, again causing severe dyslipidemia. The lipid-raising effect of clomipramine was observed both with and without co-administration of carbamazepine, implicating the tricyclic antidepressant alone. Mr. A’s levels of thyroid-stimulating hormone and fasting glucose were normal. His medical history was negative for diabetes mellitus, renal, and hepatic disease, ex-
cessive alcohol intake, and obesity. Although his father had had minor increases in total cholesterol, there was no family history of coronary heart disease, peripheral or vascular disease, or stroke among first-degree relatives.

Mr. A’s physical examination was unremarkable except for the presence of palmar xanthomas, pathognomonic of dysbetalipoproteinemia (type III hyperlipoproteinemia). He had first noticed the orange discoloration of his palmar creases 4 years after starting clomipramine but was unaware of its significance. Genetic testing confirmed APOE-2 homozygosity. Since lovastatin had previously failed to improve his lipid levels, fenofibrate, 160 mg/day, was prescribed. Unfortunately, his mood soon began to deteriorate, and he had increased frequency of suicidal ideation. Thus, although it was unclear that fenofibrate was the cause, we elected to discontinue it. Atorvastatin, 40 mg/day, was then prescribed. Subsequently, his lipid levels were somewhat improved: his total cholesterol level was 517.4 mg/dl, his triglyceride level was 460.2 mg/dl, his high-density lipoprotein cholesterol level was 57.9 mg/dl, and his total cholesterol/high-density lipoprotein cholesterol ratio was 8.9. Of interest, Mr. A complained of a depressed mood while taking atorvastatin, also necessitating its discontinuation. He recently started taking ezetimibe for lowering his lipid levels.

We suggest that the use of tricyclic antidepressants may induce dysbetalipoproteinemia in APOE-2 homozygotes. Appropriate referral and APOE genotyping should be considered for patients discovered to be dyslipidemic while taking tricyclic antidepressants.

References

Daniel T. Holmes, M.D.
Phillip Long, M.D.
Jiri Frohlich, M.D.
Vancouver, B.C., Canada

Antidepressant Effect of Ketamine During ECT

To the Editor: There has been recent interest in the use of ketamine as anesthesia for ECT because of its low anticonvulsant effects and possible reduction of cognitive side effects (1). Ketamine, a noncompetitive N-methyl-D-aspartate (NMDA) glutamate receptor antagonist, has also been shown to have putative antidepressant effects (2–4). The following case serves to highlight the possible antidepressant effects of ketamine in a patient who unintentionally received an induction dose alone during two failed ECT sessions.

Ms. A was a 47-year-old white woman who was hospitalized for an episode of severe depression in the context of a 20-year history of schizoaffective disorder. Ms. A had been treated for depression with ECT 8 years previously and had experienced profound confusion and short-term memory problems. However, she did respond, with remission of her depression. The current episode of depression had persisted for 16 weeks, during which she did not respond to venlafaxine, bupropion, sertraline, olanzapine, or lamotrigine. She had been simultaneously taking valproic acid. Ms. A’s depression was characterized by profound dysphoria, anergia, anorexia, insomnia, anhedonia, and passive suicidal ideation. There was no observed or reported mood variability, except for morning worsening of the dysphoria. She was hospitalized because of a decline in her activities of daily living to the point of not dressing herself without assistance.

Ms. A complained of severe memory problems, although her score on the Mini-Mental Status Examination was 30 of 30. The results of an EEG and magnetic resonance imaging were normal. Ms. A was withdrawn from valproic acid, 1500 mg/day, and lamotrigine, 50 mg/day, 24 hours before her first ECT treatment. For the index ECT treatment, ketamine was used as an induction agent at a dose of 0.5 mg/kg because we have found that it reduces cognitive side effects in some patients. Ms. A had no previous exposure to ketamine. Bifrontal lead placement was used, and a stimulus with the Spectrum 5000 Q ECT apparatus (MECTA Corp., Lake Oswego, Ore.) was administered by using the dose-titration method. The cuff method was used to monitor the motor seizure, and the electrical seizure was monitored with an EEG. No motor or electrical seizure was observed during the first treatment session.

Ms. A reported an immediate improvement of her mood after regaining consciousness. This improvement continued the next day when she awoke in the morning and reported a subjective improvement in well-being and an appetite for the first time in 2 weeks. The following day, ECT was again administered with the dose-titration method because we assumed that the antiepileptic agents interfered with the induction of a seizure during the first treatment session. Again, no electrical or motor seizure was observed. Strikingly, Ms. A reported a further improvement in her mood that persisted the following day. Her core symptoms, interest, energy, motivation, mood, and appetite all improved. On our 0–10-point rating scale of clinical improvement, Ms. A rated herself at 7, having initially rated herself at 2. Session 2 fell on a Friday, so there was an extra day to observe Ms. A’s response.

Forty-eight hours after her second ECT session, Ms. A still noted improvement. She did note some decline in her mood and felt that the improvement was starting to dissipate. Because of the timing of the treatments, we had 5 days to observe her clear improvement over baseline in response to ketamine. At session 3, a full grand mal seizure was observed. Three more treatment sessions were necessary before remission was reached.

Ketamine, as both an adjunctive dose (3) and an inductive dose (4), has been noted to improve mood in patients undergoing surgery. Two studies have demonstrated its potential antidepressant effects in major depression (2, 3). Of interest, acute administration of NMDA antagonists has been demonstrated to induce neurogenesis, a characteristic seen in many antidepressant agents. NMDA-mediated involvement of the glutamatergic system in the pathophysiology of depression is...
of growing interest (5, 6). Ketamine is also a weak dopamine transporter antagonist and can be psychotomimetic. Its role as a primary anesthetic agent in ECT deserves more study. It may have its greatest use in patients with severe nonpsychotic depression.

References

ROBERT OSTROFF, M.D.
MARBIELA GONZALES, M.D.
GERARD SANACORA, M.D.
Hamden, Conn.

Sertraline- and Mirtazapine-Induced Severe Neutropenia

To the Editor: Neutropenia is a rare and reversible side effect of antidepressant treatment (1–4). Six cases of agranulocytosis in approximately 2 million exposures to mirtazapine have been reported (unpublished data from Organon, Inc., 1999). Because all of the patients had either a concomitant medication or disease that might have been related to agranulocytosis, it has been suggested that the association between mirtazapine and agranulocytosis might have been coincidental (5).

We report on a person with depression who developed severe neutropenia during treatment with mirtazapine and was safely treated with sertraline.

Ms. A, a 44-year-old woman with complaints of sleep disturbance, lack of energy, and unhappiness, was diagnosed with major depressive disorder and administered mirtazapine, 30 mg/day. Her medical history was negative, and the results of routine blood tests (WBC count of 6.8 × 10^9/liter) were unremarkable except for a total cholesterol level of 204 mg/dl. Three weeks later, she came in with complaints of a severe sore throat, difficulty swallowing, a loss of appetite, and an aphthous ulcer in her oral mucosa, with a slightly elevated body temperature of 37.7°C during her physical examination. In subsequent blood counts, neutropenia (a WBC count of 2.2 × 10^9/liter) was detected, and a blood smear revealed a granulocyte number of 1.1 × 10^9/liter. Mirtazapine was immediately discontinued, and sultamicillin, 375 mg b.i.d., was started after consulting with the otorhinolaryngology department. Within 2 weeks, Ms. A's WBC count and granulocyte count had gradually increased to 3.8 × 10^9/liter and 3.2 × 10^9/liter, respectively. Four weeks after discontinuation of mirtazapine, sertraline was administered at 50 mg/day. Within 6 weeks, Ms. A's depression had responded to treatment, with more than a 50% reduction of her score on the Hamilton Depression Rating Scale, while her WBC count was 6.1 × 10^9/liter. Upon her final assessment, after 6 months of taking sertraline, her depression had remitted completely, without any adverse effects.

Since there was neither concomitant medication nor medical illness, an association between mirtazapine and severe neutropenia might be suggested. Neutropenia with cross-intolerance between two tricyclics has been described before (6), and there is evidence that patients may successfully be treated with another class of drug after such an incidence (4). Therefore, a selective serotonin reuptake inhibitor, sertraline, was administered and successfully used without an adverse effect. We may hypothesize that a different class of antidepressants might cause agranulocytosis by different mechanisms. Patients should be monitored closely for symptoms indicating agranulocytosis. An antidepressant from a different class might be considered after such an incident.

References

TUBA OZCANLI, M.D.
BARIS UNSALVER, M.D.
SAMURAY OZDEMIR, M.D.
MINE OZMEN, M.D.
Istanbul, Turkey

Clozapine-Induced Eosinophilic Colitis

To the Editor: Clozapine is an atypical antipsychotic with considerable efficacy compared to other antipsychotic medications (1). We report on a patient who developed fever and diarrhea while taking clozapine and was diagnosed with clozapine-induced eosinophilic colitis.

Mr. A was a 45-year-old man with schizophrenia who had psychotic decompensation in the setting of medication noncompliance and developed neuroleptic malignant syndrome when haloperidol and risperidone were restarted. He was treated with ECT. Concurrently, clozapine was started at a low dose and gradually increased. On the 14th day of clozapine therapy, he developed a fever of 103.6°F and profuse nonbloody diarrhea. His clozapine dose was 200 mg/day. His other medications at the time were lorazepam, aspirin, and metoprolol. His other vital signs were stable, and there was no muscle rigidity or elevation of creatine kinase to suggest recurrence of neuroleptic malignant syndrome.

His laboratory values showed mild elevation of his WBC count (to 12,300/mm^3). The results of multiple blood cul-
tures, urine cultures, a chest X-ray, stool studies, and an HIV test were negative. Mr. A continued to have fever and profuse diarrhea in the ensuing 10 days. Laboratory studies were significant for peripheral eosinophilia, with a peak of 22% and an absolute count of 2,140/mm³, as well as an elevated erythrocyte sedimentation rate of 116 mm/hour. A colonoscopy was performed on the 10th day of the fever and diarrhea. Mr. A's colon was normal upon gross examination, and random biopsies were taken. Microscopically, patchy eosinophilic infiltrate and histiocytic aggregates, focally associated with crypt destruction, were noted in the architecturally preserved colonic mucosa.

Clozapine was thought to be the culprit and discontinued. There have been reports of fever and diarrhea associated with clozapine use (2). However, the etiology of these symptoms has not been clear. Eosinophilic colitis has been suggested, although there has never been any pathologic confirmation of these findings (3). To our knowledge, this is the first report of a comprehensive investigation of the etiology with extensive laboratory studies as well as microscopic visualization of colonic pathology. We note the presence of an elevated erythrocyte sedimentation rate in this case, as is often seen in eosinophilic colitis (4). Histological examination of eosinophilic colitis usually shows patchy clusters or sheets of eosinophils in the lamina propria and crypt epithelium, with or without crypt destruction (5). The histopathology observed in our patient was consistent with eosinophilic colitis.

References

RAKESH KARMACHARYA, M.D., Ph.D.
MARI MINO, M.D.
WILLIAM F. PIRL, M.D.
Boston, Mass.

Clinical Trials During the Prodomal Stage of Schizophrenia

To the Editor: With criteria that reflect an “ultra high risk” for developing a psychotic disorder, the risk of converting to psychosis is from 30% to 60% over 1 year (1, 2). To our knowledge, the first published trial on the subject (1) randomly assigned 59 ultra-high-risk subjects to 6 months of active treatment (risperidone plus cognitive behavior therapy) or needs-based intervention. A second study, conducted by McGlashan and colleagues from Yale and collaborating sites from North Carolina, Toronto, and Calgary (2), used a randomized, double-blind, placebo-controlled design with 60 prodromal subjects to compare the efficacy of a low-dose antipsychotic with placebo in preventing or delaying the onset of psychosis (2).

We present recruitment and decision data from the Calgary site. During a 24-month recruitment period, 95 individuals were identified as potentially eligible for the study. Based on study entry criteria, 38% (N=36) were eligible; 86% of the eligible persons refused to enter the trial. Of those eligible, 39% who were troubled by their symptoms wanted active treatment as soon as possible and were offered treatment in our Early Psychosis Program; 47% of the eligible subjects were concerned, often felt debilitated by their symptoms, but did not want to take medication. This group either consented to enter a 1-year observational study offering education, support, and monthly assessments or attended a session to receive education about their symptoms. All had immediate access to a clinician should they experience an increase in symptoms in the near future. The remaining 14% consented to enter the medication trial. An assessment of presenting symptoms with the Scale of Prodromal Symptoms (2) demonstrated that those who wanted treatment and those who entered the trial had significantly higher scores on both the positive and negative symptom scales of the Scale of Prodromal Symptoms (p<0.05) than those who did not want active treatment.

Clinical trials examining the impact of antipsychotic medication in preventing or delaying schizophrenia psychoses likely recruit a small proportion of eligible individuals. The fact that those who enter such a trial have greater symptoms than those who refused to participate in the trial demonstrates that putatively prodromal individuals who enter medication trials are sick individuals with significant impairment who may potentially be on the cusp of conversion. The majority of eligible subjects who have less marked symptoms prefer to be involved to some degree in a range of educational and psychological interventions.

References
1. McGorry PD, Yung AR, Phillips Lj, Yuen HP, Francey S, Cosgrave EM, Germano D, Bravin J, McDonald T, Blair A, Adlard S, Jackson H: Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. Arch Gen Psychiatry 2002; 59:921–928

JEAN ADDINGTON, Ph.D.
DONALD ADDINGTON, M.D.
Calgary, Alberta, Canada

The Significance of Homocysteine Levels in Schizophrenia

To the Editor: We noted with interest the findings of Donald C. Goff, M.D., et al. (1) but question whether the comparison group they used was representative of homocysteine measurement. Differences in age and homocysteine measurement and an absence of dietary history and genotyping for methylenetetrahydrofolate reductase, which are important
confounding variables, may have biased the homocysteine analysis.

The study group’s mean age was 20 years younger than the comparison group’s (the sixth Framingham Offspring cohort subgroup). Homocysteine levels rise progressively with age, approximately doubling from childhood to old age. There is a bias toward higher comparison homocysteine levels as a result (2, 3).

The accuracy of homocysteine differs among methods and laboratories. This effect had not been controlled. The Framingham Offspring comparison group’s homocysteine level was measured with high-performance liquid chromatography with fluorescent detection, whereas the study group used fluorescence polarization immunoassay. The use of a local comparison group with the same method and laboratory is recommended (3).

The Framingham Offspring Study shows that mandatory fortification policy has reduced the prevalence of low folate status (i.e., <3.0 ng/mL) by more than 90% and the prevalence of mild fasting hyperhomocysteinemia (homocysteine concentrations >130 mmol/liter) by about 50% among its population-based cohort of middle-age to elderly U.S. citizens (4).

We recommend that the assessment of homocysteine requires the measurement of known confounding variables, especially in smaller group sizes, from folic-acid-fortified regions. Contrary to larger previous studies with local comparison groups showing increased homocysteine levels in schizophrenia, these results should be treated with caution (5, 6).

References

COLIN O’DONNELL, M.B.B.CH., M.R.C.PSYCH.
TIM STEPHENS, B.Sc.
Melbourne, Victoria, Australia

Dr. Goff and Colleagues Reply
TO THE EDITOR: We appreciate the letter from Dr. O’Donnell and Mr. Stephens and agree that our study did not clarify the question of whether homocysteine levels are elevated in schizophrenia patients. This was not the intention of our study. Our hypothesis, that serum folate concentrations would correlate inversely with the severity of negative symptoms, was based on the dual roles of glutamate carboxypeptidase II as a modulator of brain N-methyl-D-aspartic acid activity and as a facilitator of folate absorption in the intestines. This primary hypothesis was supported by our results. Although a comparison of serum concentrations with a nonpsychiatric study group was not an objective of our original study, we provided a comparison with the Framingham Offspring Study cohort as a frame of reference for our findings. Dr. O’Donnell and Mr. Stephens are correct in stating that this comparison group was not matched for age (the mean age of our group was 43 years versus 56 years for the Framingham group), nor were the same assay methods used. An ideal comparison group would be matched by age, smoking status, gender, exercise level, and diet—a daunting task. We chose the Framingham Offspring Study sample because the subjects lived in the same geographical region, were sampled after folate supplementation of grain products had been implemented, and were not taking vitamins with folate. The substantially lower folate concentrations both in our total sample and in schizophrenia nonsmokers compared to this nonpsychiatric group support the hypothesis that low folate concentrations might result from the low activity of glutamate carboxypeptidase II. As we emphasized in our discussion, there are several alternative explanations for our findings.

We did not design our study to test the hypothesis that homocysteine concentrations are elevated in schizophrenia patients compared to a healthy population, and our results should not be interpreted as such. We failed to find hypothesized clinical correlations with homocysteine serum concentrations except for a positive correlation with extrapyramidal symptoms. Age did not predict homocysteine concentrations in our group, whereas gender was a very strong predictor that should be controlled in future comparison studies. Homocysteine levels have varied widely among schizophrenia groups, having been found to be dramatically elevated primarily in young men in two groups from Israel (Applebaum et al., 2004, and Levine et al., 2002) and not elevated in groups of schizophrenia patients from the Netherlands (1) and Spain (2) and in female schizophrenia patients from Germany (3). We agree that careful examination of environmental and dietary factors will be important in understanding these findings. Clarification of the possible neuropathological role of elevated homocysteine, as found by Levine and colleagues (2002), will be of great interest to the field.

References
3. Reif A, Schneider MF, Kamolz S, Puhlmann B: Homocysteine in psychiatric disorders: association with dementia and
Reduced Spinophilin in Schizophrenia

TO THE EDITOR: Amanda J. Law, Ph.D., et al. (1) reported decreased mRNA for spinophilin in the hippocampal formation in patients with schizophrenia and mood disorders. The potential implications of this abnormality are curious since spinophilin is involved in limiting or decreasing the length of dendritic spines. Spinophilin immunoreactivity is predominantly (although not exclusively) localized to dendritic spines (2), for which it is a meaningful marker, with hippocampal spine density and spinophilin immunoreactivity both increasing in response to estrogen (3). However, spinophilin is not necessary for the formation of spines, which are increased in density in young spinophilin knockout mice or for their removal, since spine density returns to normal when such animals mature (4).

Spinophilin links protein phosphatase-1 to actin. Dephosphorylation of actin by protein phosphatase-1 leads to its disassembly (5), so the targeting of protein phosphatase-1 to actin is likely to result in fewer or shorter dendritic spines. Linking protein phosphatase-1 to actin also promotes dephosphorylation of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors (4). The subsequent decrease in the activity of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors should decrease calcium influx into the spine, reducing the phosphorylation of spinophilin by calcium-calmodulin kinase II and thereby enhancing its affinity for actin (6).

The reported subicular decrease in spinophilin mRNA in schizophrenia is approximately 30% (1). We observed a decrease of 75%–80% in the density of apical dendritic spines (plus a more modest decrease in apical dendritic arborization) (7), but the density of spines on basilar dendrites has not been measured. If the diminished pool of mRNA is supporting the synthesis of spinophilin for an even more diminished set of dendritic spines, increased synthesis of spinophilin, on a per-spine basis, might function to keep spine density reduced. Posttranscriptional mechanisms are also likely, as evidenced by estrogen-induced increases in spine density and spinophilin immunoreactivity, without altered levels of spinophilin mRNA (3).

References


ANDREW J. DWORK, M.D.
GORAZD ROSOKLIJA, M.D., Ph.D.
LIESL B. JONES, Ph.D.
New York, N.Y.

Dr. Law and Colleagues Reply

TO THE EDITOR: Dr. Dwork and colleagues raise several interesting issues. We emphasized that our main conclusion was merely that the decrease in spinophilin mRNA provides a further indication that the site of molecular alterations in synapses in schizophrenia and mood disorders includes dendrites as well as presynaptic terminals. We also suggested, parsimoniously, that the reduction is probably related to the decreases in spine density, which the correspondents and others have clearly demonstrated. Given what is known of the functions of spinophilin, as outlined in their letter, we agree that the reduction in its mRNA is more likely to be a consequence than a cause of the lower spine density (i.e., fewer spines require synthesis of less spinophilin), as we pointed out in our article (p. 1862). The fact that the percentage decrement in spine density is much greater than the reduction of spinophilin mRNA can be explained, as Dr. Dwork and colleagues note, by the fact that changes in transcript abundance are often less than those of the encoded protein and that the subcellular difference in the targeting of proteins may exist. Along similar lines, the lack of change in microtubule-associated protein 2 mRNA that we find may signify that any putative changes in microtubule-associated protein 2 are translational or posttranslational in origin.

Dr. Dwork and colleagues proposed one cascade in which spinophilin is involved and regulated. However, this important basic work was largely derived from model systems, and the extent to which these molecular cascades are operating in the predicted manner in human hippocampal neurons is not known.

In summary, we did not advocate that spinophilin is a molecule with a specific or central etiological role in schizophrenia or mood disorder, and we were unaware of the complexity of spinophilin’s roles in dendritic function. Rather, as the title of the article stated, the aim was to complement and draw attention to the evidence that dendritic spines are part of the molecular and cellular neuropathology of schizophrenia and mood disorder—work at which Dr. Dwork and colleagues have been at the forefront. We are sure that they would agree that more studies are necessary to refine the evidence for
dendritic (spine) pathology and to begin to reveal the biochemical pathways that underlie it.

AMANDA J. LAW, Ph.D.
CYNTHIA SHANNON WEICKERT, Ph.D.
THOMAS M. HYDE, M.D., Ph.D.
JOEL E. KLEINMAN, M.D., Ph.D.
PAUL J. HARRISON, D.M., F.R.C.PSCH.
Oxford, U.K.

Physician Suicide and Drug Abuse

TO THE EDITOR: Physician suicide rates and suggestions for future studies were nicely reviewed by Eva S. Schernhammer, M.D., Dr.P.H., and Graham A. Colditz, M.D., D.P.H. (1). However, they might consider further evaluation of the “risk factors relating to the working environment” (p. 2300). We have reported on outcomes of impaired physicians for nearly 25 years and have followed all impaired Florida physicians since 1995 (2, 3). Physician drug abuse has been linked to suicide (4, 5). We suggested that not all physician specialties are equally affected by drug abuse and dependence. Similarly, suicide may affect one medical specialty more than another. We have suggested workplace evaluations, starting with a history of drug exposure in the operating and emergency rooms and intensive care units. Anesthesiologists are significantly overrepresented among Florida physicians with substance use disorders. They represent only 5.6% of the total licensed physicians but almost 25% of the physicians with substance use disorders. Access to drugs of abuse has been the major theory advanced to explain this. However, we have proposed that unintended second-hand environmental exposure puts anesthesiologists at increased risk (6). We also recently demonstrated the presence of propofol and fentanyl in operating room air after intravenous administration (7). Secondhand exposure is a fact in some medical workplaces. It was not surprising that anesthesiologists and other physicians exposed to fentanyl in the workplace represented 90% of the fentanyl abusers in Florida. Studies in progress include sampling anesthesiologists’ blood during work in cardiovascular surgery, an environment where high doses of fentanyl are routinely used. Environmental exposure may explain the high rates of addiction among anesthesiologists and why recovery for anesthesiologists often necessitates giving up their work in operating rooms and even changing medical specialties. Prevention of physician opioid abuse and dependence appears to be linked to identifying sources of secondhand exposure and preventing exposure from occurring or by minimizing exposure, as was done with nitrous oxide. Environmental exposure may also prove to be an important factor in suicide attempts, relapses, and prevention. We would strongly suggest that new and important data from the analysis of Drs. Schernhammer and Colditz be expanded to include medical subspecialty and secondhand exposure.

References

MARK S. GOLD, M.D.
KIMBERLY FROST-PINEDA, M.P.H.
RICHARD J. MELKER, M.D., Ph.D.
Gainesville, Fla.

Dr. Schernhammer Replies

TO THE EDITOR: My associates and I appreciate the additional background information conveyed by Dr. Gold and colleagues. Although presenting data on physician suicide rates by medical specialty was beyond the scope of our analysis, we undoubtedly welcome their invitation for more research.

Generally, men tend to successfully commit suicide more often than women, whereas suicide attempts among women are higher than those of men. It is possible, however, that women physicians more successfully commit suicide than do women outside of the medical profession. Among U.S. medical students, the observed suicide rate of female students equaled that of the male students (although they were still three to four times higher than those of their age mates), indicating a relative scarcity of attempted suicides in that profession (1).

That access to drugs can support higher suicide rates has long been shown: for example, in Australia, an increase in suicides by women coincided with the implementation of a law that facilitated access to barbiturates (2). Therefore, it appears likely that higher suicide rates among physicians, who tend to prefer methods that are typical for their profession (3), may be coupled with both their easier access to drugs, as well as their better know-how concerning the successful use of such methods. The interesting proposal by Dr. Gold et al. of unintended secondhand environmental exposure to drugs as a risk factor for drug addiction and possibly suicide among anesthesiologists adds another layer of complexity and warrants further investigation.

References

EVA SCHERNHAMMER, M.D., Dr.P.H.
Oxford, U.K.
Methodological Concerns in a Trial of Ziprasidone and Olanzapine

TO THE EDITOR: In their randomized, double-blind trial comparing ziprasidone and olanzapine for the treatment of acutely ill inpatients with schizophrenia or schizoaffective disorder, George M. Simpson, M.D., et al. (1) provided important information showing that olanzapine-treated patients have a greater risk of weight gain and lipid abnormalities than patients treated with ziprasidone. However, the dosing protocol in this study raised a number of questions. First, there appeared to be a potential for unblinding. Each blister pack of study medication was labeled “A,” “B,” or “C,” corresponding to a “low,” “medium,” or “high” dose of each drug. All ziprasidone-treated patients were to receive the “high” dose at the end of 1 week, whereas the olanzapine-treated patients received the “medium” dose. During the trial, the treating clinician would need to know the current dose classification each week to decide whether it should or could be increased or decreased. A “medium” dose after the end of the first week would clearly indicate olanzapine treatment, whereas a “high” dose would indicate ziprasidone treatment. It is possible that unpublished procedures were used to prevent this potential problem. If so, knowledge of these procedures would be helpful in interpreting the results of the trial.

A second concern with regard to the dosing protocol is one that is not uncommon in trials sponsored by pharmaceutical companies, that of a suboptimal dose of a comparator drug. In this trial, the patients could receive a maximum olanzapine dose of only 15 mg/day, although the product labeling recommended doses up to 20 mg/day (2). The patients received 10 mg/day at the end of 1 week, and the mean dose of olanzapine throughout the trial was only 11.3 mg/day. In contrast, ziprasidone was titrated to the maximum dose recommended by the product labeling (3), 160 mg/day, by the third day of the trial. In order to reduce bias in studies comparing drugs of a sponsor and competitor, available doses should include the entire range recommended by the product labeling.

References

RYAN M. CARNAHAN, PHARM.D., M.S.
Tulsa, Okla.

PAUL J. PERRY, PH.D.
Iowa City, Iowa

Dr. Simpson Replies

TO THE EDITOR: I thank Drs. Carnahan, Perry, and Ross for their comments on our study and welcome the opportunity to discuss the issues they raise regarding drug administration, concommitant medication, and dosing.

Drs. Carnahan and Perry comment that the medication labeling—“A,” “B,” and “C,” denoting “low,” “medium,” and “high” doses—would alert investigators at day 7 to treatment assignment because all patients randomly assigned to ziprasidone received 80 mg b.i.d. (the “high” dose) on days 3 to 7 and all patients randomly assigned to olanzapine received 10 mg/day (the “medium” dose). The “A,” “B,” and “C” labeling was, in fact, employed only during the flexible-dose weeks of the study (weeks 2 to 6). During days 1 to 2 and days 3 to 7, when fixed doses were administered, the medication cards did not contain this labeling. During both the fixed titration and flexible-dose phases of the study, the patients in the two treatment arms received identical quantities of medication of identical appearance; in the olanzapine group, placebo capsules were employed to simulate twice-a-day dosing. For example, a subject who was assigned to olanzapine at 10 mg/day would have received the same number of identical-appearing capsules twice a day as a subject who was assigned to ziprasidone at 80 mg b.i.d. with this “double-dummy” design.
Thus, there was no potential for unblinding in the drug-administration protocol.

Dr. Ross questions the use of lorazepam in our trial to treat agitation or insomnia. Clinical trials of antipsychotics in acute schizophrenia have commonly permitted concomitant use of benzodiazepines, although data on their use have not always been reported. The percentages we reported in our article for any use of lorazepam during the study—83.1% in the ziprasidone group and 75.2% in the olanzapine group—reflected lorazepam use during days 1 to 7. During days 8 to 14, the percentages of subjects still taking lorazepam decreased to 51.5% in the ziprasidone group and 51.9% in the olanzapine group. Lorazepam use continued to decrease over subsequent study weeks and remained comparable for the two treatment groups throughout the study.

With regard to the comments on dosing, the olanzapine dosing regimen employed in the trial was consistent with dosing recommendations from the olanzapine prescribing information, as well as with published clinical trial data available at the time the study was designed and conducted (1–4). It should be noted that only olanzapine-naïve patients were included in our study. Patients who had more than 14 days of total lifetime exposure to olanzapine or who had received an olanzapine dose above 10 mg/day were excluded. This criterion for entry, together with the fact that patients who are required to provide informed consent in such clinical trials are generally not of the highest level of illness severity, may have moderated the average dosing requirements for subjects in the study. The mean olanzapine dose in our study reached 13 mg/day by days 15 to 21 and remained at that level for the remaining study weeks (the mean dose for the entire length of the study was 11.3 mg/day), indicating that this medication was adequately dosed. As noted in our article, the mean olanzapine dose after week 5 in our study—13.1 mg/day—is virtually identical to the endpoint mean dose in the comparative trial of olanzapine and risperidone reported by Conley and Mahmoud (5).

Although ziprasidone could be administered at up to 160 mg/day, the maximum dose currently recommended in the product labeling (the package insert for Geodon), the mean doses were approximately 140 mg/day at days 15 to 21 and the following weeks, with a mean dose of 130 mg/day for the entire length of the study, indicating that many subjects were judged not to require the maximum recommended dose. The percentages of subjects receiving the maximum dose at endpoint were similar for the olanzapine (59.4%) and ziprasidone (61.8%) groups. In addition to observing comparable efficacy between ziprasidone and olanzapine in the treatment of hospitalized patients with acute schizophrenia, we found differences between treatment groups in body weight, lipid profile, and insulin resistance. Presumably, higher doses would not have ameliorated—and may have exacerbated—these adverse events observed in the olanzapine-treated subjects.

References

GEORGE M. SIMPSON, M.D.
Los Angeles, Calif.

Olanzapine and Haloperidol for Residual Symptoms

TO THE EDITOR: Robert W. Buchanan, M.D., and colleagues (1) conducted an important study on the comparative effect of olanzapine and haloperidol on residual positive and negative symptoms in 63 outpatient patients with treatment-resistant schizophrenia. They concluded that olanzapine has limited differential benefit for positive and negative symptoms in these patients. However, we think that the correct conclusion must be that olanzapine has no benefit over haloperidol. Namely, the authors could not find superior efficacy for olanzapine for positive and negative symptoms, and the main side effects (extrapyramidal symptoms for haloperidol and weight gain for olanzapine) seem to balance each other out. The authors stated that “the magnitude of weight gain...may potentially offset the more benign extrapyramidal symptom profile of olanzapine” (p. 128). Furthermore, we would like to point out that the mean doses of olanzapine (20.3 mg/day) and haloperidol (18.3 mg/day) resulted in noncomparable dopamine D2 receptor occupancy. For almost all patients receiving a dose of about 18 mg/day of haloperidol, the balance between efficacy and extrapyramidal side effects is not optimal (2, 3). Haloperidol at lower doses is thought to induce fewer extrapyramidal side effects and fewer neuroleptic-induced negative symptoms and dysphoria (4) without a change in efficacy on positive symptoms (2). Lower haloperidol doses might even further diminish the benefit of olanzapine for extrapyramidal symptoms.

Therefore, the sobering conclusion of this study seems to be that neither olanzapine nor haloperidol shows benefits for outpatients with schizophrenia who met criteria for either residual positive or residual negative symptoms.

References

LIEUWE DE HAAN, M.D., PH.D.
Amsterdam, the Netherlands
NICO VAN BEVEREN, M.D.
Rotterdam, the Netherlands

Dr. Buchanan and Colleagues Reply

To the Editor: Drs. de Haan and van Beveren raise two points concerning our recent article on the comparative efficacy of olanzapine and haloperidol for residual positive or negative symptoms. First, they note that in the abstract of our article we stated that “Olanzapine has limited differential benefit for either positive or negative symptoms in patients with treatment-resistant schizophrenia” (p. 124). They suggest that the proper conclusion should be that “olanzapine has no benefit over haloperidol.” The qualification in our statement reflects the fact that our study is not the only one to have addressed this issue. In our Discussion section, we reviewed two studies that asserted a benefit for olanzapine in this population (1, 2). Although we believe that these studies have methodological flaws, they remain in contradistinction to our results.

Second, Drs. de Haan and van Beveren suggest that the mean doses of haloperidol and olanzapine achieved in our study result in noncomparable D₂ receptor occupancy, with the haloperidol dose associated with increased D₂ receptor occupancy, which could potentially lead to increased extrapyramidal symptoms, dysphoria, and secondary negative symptoms. Although we agree with the theoretical concern, we note that the haloperidol-treated patients did not exhibit a mean worsening of either extrapyramidal symptoms or depressive symptoms (as measured by the Hamilton Depression Scale) nor a worsening of negative symptoms (Tables 2 and 3). The alternative concern, which we raised in our Discussion section, is that the olanzapine dose was relatively low for this population. A higher olanzapine dose may have led to increased symptomatic improvement.

References

ROBERT W. BUCHANAN, M.D.
M. PATRICIA BALL, R.N.
ELAINE WEINER, M.D.
Baltimore, Md.

Origin of the Term “Schizophrenia”

To the Editor: The portrayal by Ernest L. Abel, Ph.D., of the 1857 “theory of degeneration” of Benedict-Augustin Morel (1) as a “parsimonious explanation for the etiology of insanity” dominating French psychiatry for “almost a century” overestimated its influence by almost half a century.

By 1911, Swiss psychiatrist Paul Eugen Bleuler had renamed Kraepelin’s 1889 Latin form of Morel’s earlier term demence precoce, “schizophrenia,” emphasizing that the illness known as “dementia praecox” was not an actual dementia and did not always begin at an early age. Although the growing influence of German psychiatry in France induced negative reactions, whose main target was the work of Kraepelin, the nationalistic tone reached its apex on the French side during World War I, and only some faint traces remained for several years after 1920 (2).

Moreover, the 1950 discovery of chlorpromazine by physician Henri Laborit and the seminal work in 1952 by French psychiatrists Jean Delay and Pierre Deniker, which introduced the chemical as a treatment for schizophrenia (3), demonstrated French psychiatrists’ much earlier acceptance of the Bleuler model.

Although today we can look to Morel’s work as a progenitor of the current biological approach to psychiatric illnesses, Bleuler’s characterization—the group of schizophrenias—remains the more parsimonious approach.

References
2. Pichot P: [The history of German psychiatry from the viewpoint of French psychiatrists]. Fortschr Neurol Psychiatr 1992; 60:317–328 (German)

JACK C. SCHOENHOLTZ, M.D.
Mamaroneck, N.Y.
The Neurological Basis of Pain provides the latest thinking and practice in the field and helps the reader grasp much of the vast biology of neuroplasticity, the basis of neuropathic pain. Focusing on symptoms and etiology has not provided a method of intervention, however; similar symptoms may arise from different etiologies, and the same etiology may be associated with a range of symptoms. Thus, research is progressing to describe the underlying mechanisms and the means by which they can be clinically identified and managed. Dr. Portenoy, in the foreword, states that the development of new drugs and other treatments “holds the promise of true mechanism-based therapy.” Thus, the wheel may turn again and the biopsychosocial approach may become redundant.

There is a dearth of psychiatrists among the authors, but psychologists are well represented. This reflects real-world pain management, although psychiatrists have much to offer in this field, and the work is rewarding. In a chapter on spinal pain, Wheeler and Murrey state that training should be made available to neurologists and psychiatrists so they can establish “competence in using electrodiagnostic skills for diagnosis and performing spinal interventional treatments with spinal injections.” I hope this comes to pass.

The Neurological Basis of Pain is a magnificent contribution to pain management. It has the combination of readability and scholarship that will make it a classic. It should be in every medical library and on the shelf of everyone with a serious interest in pain medicine.

References

SAXBY PRIDMORE, M.D.
Hobart, Tasmania, Australia


This multi-authored book provides a practical overview of clinical neurology for physicians who work in other areas of medicine. It begins with a concise and well-focused discussion of the basic neurological examination, beginning with the cranial nerves and peripheral nervous system and moving systematically through basic diagnostic tests and the evaluation of coma. Each chapter is individually authored and covers a broad variety of neurological problems, including cerebrovascular disease, headache, epilepsy, multiple sclerosis, Parkinson’s disease, alcoholism, peripheral neuropathy, the dementias, traumatic brain injury, neuromuscular disorders, and neurological manifestations of cancer. There is also an entire chapter devoted to neurological complications of pregnancy. These topics are all dealt with in a straightforward way that is suitable for the average psychiatrist. NeuroAIDS and other conditions with relevance to psychiatry, such as low back pain and insomnia, are also dealt with in a clear and practical way.

Most chapters provide a description of the correct ways to perform a neurological examination and enumerate the...
forms of treatment that are available for different neurological disorders. In the chapter on “eye signs,” detailed diagrams showing visual pathways are provided and visual field defects are depicted in simple diagrams. The final chapter includes a discussion of medical-legal issues involved in the care of patients with neurological disorders as well as a discussion of the problems associated with obtaining informed consent from patients who are not legally competent. This chapter also contains a fairly extensive discussion of brain death, including the use of electroencephalography and cerebral blood flow as part of the evaluation. For the psychiatrist who is seeking to learn more about behavioral neurology, there is a description of the Papez circuit and a fairly comprehensive discussion of temporal lobe epilepsy. Beyond this, however, there is very little information regarding psychiatric manifestations of neurological disorders. For those who treat the geriatric population, where behavioral manifestations of delirium and dementia are common, this book is probably not an appropriate resource. Overall, however, Neurology for the Non-Neurologist provides a well-illustrated and clearly written overview of basic clinical neurology.

Francine M. Benes, M.D., Ph.D.
Belmont, Mass.


I was hopelessly bad at mathematics at school and was happy to leave a world of half-understood quadratic equations and integral calculus well and truly behind me in my teens. Yet, a teacher’s end-of-term report on my progress in the subject has stuck with me as a cutting, if accurate, description of more generalized ability: “Tries hard—but then he has to.” This reassured my parents that I was at least doing my best and gave me a useful insight into my personal limitations. Many years older and not much wiser, I appreciate how much harder it is to write an honest report than one in which positive gloss, thickly applied, obscures the message.

In their introduction, the editors tell us that this volume in the Review of Psychiatry series provides an overview of the neuropsychiatric approach to assessment and presents discussions of techniques and testing methods that may be more familiar to neurologists than to psychiatrists. The book certainly does what it claims to on the cover: five chapters review the neuropsychiatric and neuropsychological examinations, electrophysiological and pathological laboratory testing, and neuroimaging.

The book is attractively produced, and when I had finished reading the introduction to the Review of Psychiatry series that occupied the first few pages, I feel excited by the prospect of what lay ahead. I am very sorry to say that I was disappointed by all five of the book’s main chapters. None of them offers much more than a beginner’s guide to the subject, and some are woefully patchy in their coverage of chosen areas. The more I read, the more puzzled I became as to exactly who the target readership for this book might be. Medical students or perhaps the most junior of psychiatric or neurological trainees will find competent accounts of neuropsychological and neuropsychiatric assessments here, but any good core textbook would already cover these areas and more besides. More senior colleagues will not find the definitive, up-to-date reviews that they would otherwise look for in bigger textbooks or journals. I could not think of a particular group of psychiatrists, psychologists, or neurologists to whom I would recommend this book. There will be future titles in this series, and my old teacher would perhaps have written, “We are hoping to see some of the promised talent in the next term.”

Robert Howard, M.D., M.R.C.Psych.
London, U.K.

Sleep Disorders


The study and treatment of sleep disorders is a young, rapidly expanding specialty, with board certification, accredited treatment centers, and research journals. Combining aspects of psychiatry, neurology, and pulmonary medicine, it deals with an important human function largely ignored until the discovery of rapid eye movement (REM) sleep (1). Half a century later, hundreds of sleep laboratories offer diagnostic studies, and scores of accredited sleep disorder centers provide both diagnosis and ongoing treatment for insomnia, obstructive sleep apnea, narcolepsy, and other sleep-related illnesses.

The authors’ stated audience for this compact volume includes trainees, practicing sleep physicians, and academic training programs, and they hope to spur interest in those outside the field. They reflect their own practice at the Mayo Clinic: for example, they describe a night in their sleep laboratory that differs in some interesting details from the procedures of other accredited centers. Their overview is organized into four sections: Basics of Sleep Medicine, The Sleepy Patient, The Patient Who Cannot Sleep, and The Patient With Excessive Movement in Sleep. In this general approach lie both the strengths of this book and its weakness.

As a short textbook, Sleep Medicine in Clinical Practice is a useful addition to the available general literature, suitable both as an overall introduction and as a quick reference for busy practitioners. It covers the diagnostic categories within the current (soon to be revised) International Classification of Sleep Disorders, some very briefly. Clinical norms, often adjusted by age group, will help the general physician interpret sleep study reports. Illustrative clinical vignettes are well placed and helpful, fleshing out the diagnostic entities. Charts, tables, and line drawings illustrate key topics and provide examples of polysomnographic studies. Useful clinical algorithms show how to reach clinical diagnoses and treatment decisions.

The authors’ effort to provide a wide overview of sleep medicine, however, ultimately limits the book’s usefulness. Rare and unusual diagnoses are covered, but the same succinct approach is applied to common disorders. For example, the authors describe recurrent (periodic) hypersomnia and catathrenia (nocturnal expiratory groaning), illnesses the average
organization, and mental disorders can be seen as perturbations of this organization. For example, Peled sees dysthymia as “current deoptimization shifts of the transmodal levels accompanied by constraint frustration” (p. 100), bipolar mood swings as an “oscillatory dynamic of optimizations and deoptimizations” (p. 72), psychosis as “connectivity breakdown of the dynamic core” (p. 72), anxious loss of control as “destabilization of the higher level transmodal brain systems relevant to conscious awareness” (p. 73), and transference as “activation of the attractor systems which represent the person from the past” (p. 75). He does not address character types, which I believe can be modeled plausibly by tweaking neural network elements.

Peled provides illustrative cases for his system of psychiatric brain profiling in which both “external and internal perturbators affect the system development and organization” (p. 93). For external perturbators he favors the Holmes-Rahe Social Adjustment Scale (p. 94), which ranges from minor violations of the law and Christmas to divorce and death of spouse. Internal perturbators include metabolic, medication, and intracranial pathological effects. Peled says his Psychiatric Brain Profile is less stigmatizing and categorical and has more “degrees of freedom” than DSM and yet is more constrained by neuroscience than psychoanalytic conceptualizations that “have unlimited degrees of freedom allowing for all concepts to describe all occurrences and thus are operationally meaningless” (p. 95). Many of Peled’s sources as well as his subtitle, however, are conceptually derived from psychoanalysis.

The book concludes with ideas about future directions for psychiatry. Testable hypotheses must move from a linear two-factor model of cause and effect to a three-level model of lower-level multiple biological factors, an intermediate-level system organization, and higher-level system functions and emergent functions. Means to detect perturbations of brain organization must be developed. Synaptogenic control should include neurogenesis. Direct pacemaker control could include transcranial magnetic stimulation coupled with EEG as well as imaging and deep brain stimulation. Experience control should include virtual reality technology, which Peled believes has potential to correct specific brain cognitive deficiencies, even delusions (for example, by showing patients with delusions of persecution by the FBI a warm and caring FBI headquarters).

If the brain were a corporate office, our present state of functional imaging would put us in the position of the superintendent in the basement who can monitor departments’ use of electrical power and tell who is burning the midnight oil. Peled seems to be proposing that psychiatrists assume the role of a corporate information technology manager who knows the information storage and transmission requirements of each department as well as the volume and destinations of its e-mail and who addresses bugs, overloads, crashes, and viruses that arise in the system. In a continuation of the metaphor, psychoanalysis would be a little like entering the play sphere of the office party to observe the employees’ interactions, or perhaps taking the chief executive officer’s secretary to lunch to hear gossip about office politics. We still do not know how to read the e-mail and must infer how the corporation does its business, decides its priorities and strategies, innovates, integrates its employees’ expertise, and sets departmental budgets.


Given the expansive title of this book, I expected to find a cryptic and complex tome of many hundreds of pages filled with diagrams, positron emission tomography scans, and the most modern three-dimensional images. As it turned out, I was partially correct. The Mind does seem cryptic and complex to me, but it runs under 250 smallish pages and fulfills the promise implied about the nature or origin of one’s mind.

I am willing to accept some responsibility for not fully grasping the merits of the book. Perhaps I am too narrowly focused in biologically oriented psychiatric medicine to appreciate this effort. However, some fault must lie with the author and editors for producing a book whose book jacket summary, a reflection of the contents, is so convoluted that it left me and a few colleagues puzzled.

The quest, we are told, is to demonstrate how cerebral activities become mental events. On this journey of 21 chapters organized within five parts we revisit Freud, Piaget, and many other revered psychologists and behavioral scientists; engrams; and Gestalt. We are treated to the history of dialectical concepts, beginning with Zeno of Elea in 464 BCE. We are offered dialectical interconnectedness and dialectical triads. We are awash with matrices and fusions as well as complex and simple mnemonics. Part 3, Structures of the Mind, has a chapter titled “Ego, Superego, Id,” and the only other chapter is titled “The Neurophysiology of Dreaming.”

The section on Brain, Mind and Body includes six chapters totaling 28 pages. The chapter on the “Mind-Body Problem” is less than four pages long and has three references, from 1950, 1980, and 1985 (by the author), and includes a figure (number 10) of the ubiquitous smiley face and sad face.

The section on Psychological and Clinical Implications covers personality formation, psychopathology, and psychotherapy in a brief 40 pages. Although it might fortify the previous chapters, as a clinician I found that it did not offer much to strengthen my practice or pearls to share with residents.

Pointing out what I see as flaws in this book does not necessarily mean I found no merit. I look at The Mind as primarily a relatively brief philosophical, psychological, and scientific exploration of concepts. For those wanting to ponder “the mind,” memory, and cognition relative to a historical framework it may well be a very rewarding experience. For those looking for a more useful clinical or teaching tool for psychiatric medicine I do not think this would be a first-line choice.

MARK H. FLEISHER, M.D.
Omaha, Neb.


The 30 chapters of this book are written by 25 philosophers and a handful of clinicians. Most are written by a single author, addressing the many philosophical aspects of psychiatric practice and targeting issues such as what it means to make a diagnosis, when is one not responsible for a criminal act, what does competence imply, and when is the self continuous or discontinuous. There are philosophical discussions of such varied topics as brain pain, desire, memory, values, evolution, research ethics, religion, race, and gender.

I wish I could say that the promise inherent in the nature of the topics translates into great reading, but it doesn’t. The idea for the book is excellent, and this may prove a very successful text if the intended readers are philosophy students. Clinicians, however, will find most of the chapters hard to read. The language of academic philosophy is not the language of bedside psychiatry. With some exceptions, the single-authored chapters by philosophers are so dense, so laden with jargon, and so embedded in a philosophical context inscrutable to the ordinary reader that their message is lost.

A notable exception is the chapter by Jennifer Church on the social construction of madness. This is a wonderful exploration of the pros and cons of viewing illness as socially constructed versus accepting it as biologically predetermined. This chapter is insightful, thought-provoking, and written in plain English. Another very readable, balanced, and useful chapter written by a philosopher, Daniel Robinson, is on the concept of dangerousness. On the whole, the chapters that work best are those written by a philosophy/psychiatry team. Sadly, there aren’t many of these. Michael Schwartz and Osborne Wiggins contribute a very good chapter dispelling the myth that clinical drug trials and neuroscience constitute the sole scientific methodologies of psychiatry. They talk about understanding and interpretation, the methodology of studying psychopathology and psychotherapy. This is important because it broadens the focus of what psychiatric practice encompasses and shields it from allegations of reductionism, charges to which biomedicine is vulnerable. Most of the co-authored chapters are worth reading—a comment on the fact that when two disciplines work together to produce a piece of writing, they abandon the jargon of their respective fields and write in a style that others can understand.

Because the idea for a book of this kind is so good, I would encourage the editor to try again. My prescription for excellence would be to restrict the number of topics and to select authors carefully. I would insist that philosophers and psychiatrists collaborate on each chapter. I would not allow bland reviews of a topic area and would insist that each chapter defend a point of view. I would ensure that all chapters be read by all authors and that wrestling with each other’s arguments be part of the task of writing. I would not permit the use of vocabulary that the general educated public does not understand. I look forward excitedly to such a book.

MARY V. SEEMAN, M.D., F.R.C.P.C.
Toronto, Ont., Canada

Am J Psychiatry 162:7, July 2005

http://ajp.psychiatryonline.org

1397

It is certainly remarkable that we still need to be reminded of the most basic, commonsense perspective that informs both the art and science of the practice of clinical medicine. Even Hippocrates understood the importance of approaching the patient in his or her social context and treating the whole person, and early in the last century Adolf Meyer was teaching a psychosocial approach to psychiatry. Our growing sophistication as trained physicians in identifying biological underpinnings of diseases should not supplant our understanding of this approach, yet in an atmosphere of more knowledge and more paperwork and less time and less continuity of care with patients, it seems we need reorientation to what our full field of vision and practice should be.

This volume celebrates the biopsychosocial approach, so aptly named and so inspiringly taught by Dr. George Engel of Rochester, N.Y. Sadly, Dr. Engel passed away in 1999, while this book was still in preparation. The relevance of his antidote to biomedical reductionism has not diminished since its introduction many decades ago. Dr. Engel teamed with Dr. John Romano, Chair of the Department of Psychiatry at University of Rochester School of Medicine, to revolutionize the medical educational culture there, and their seminal influence in that city continues to this day. Therefore, it is not surprising that 60% of the chapter authors in this volume are affiliated with the University of Rochester School of Medicine or the University of Rochester.

The biopsychosocial approach formally saw the light of day in print in Dr. Engel’s landmark 1977 Science paper (1) (reproduced in this volume), but he is more remembered in psychiatry for his 1979 APA Vestermark Award lecture, which was published in the American Journal of Psychiatry in 1980 (2). In the latter, Dr. Engel used the example of a man experiencing a myocardial infarction and suffering cardiac arrest during his medical workup. Seeing this classic reprinted as this book’s first chapter is a delight. The following chapter by the three editors, introducing the remaining chapters, delivers some unintended irony, as the three authors misremember just-presented details of that famous vignette. (For the record, the patient endured repeated unsuccessful attempts by several house officers trying to perform an arterial puncture, not start an intravenous line; and the patient said that he consequently anticipated further painful fumbling attempts and felt outrage, then self-blame and impotence, not fear of being alone and abandoned.)

“Clinical Practice and the Biopsychosocial Approach” is an impressive chapter seamlessly written by six authors, nicely bringing Dr. Engel’s description of biopsychosocial understanding to an up-to-date approach to therapeutic practices. Subsequent chapters address other clinical applications, research, educational and administrative perspectives, and, finally, some historical background and some brief notes on comparisons with systems theory and current patient-centered and relationship-centered approaches to health care.

The book is a pleasant and interesting read for those with an interest in this area, although it is written about nonpsychiatric primary care. It certainly would be relevant in training medical students and residents, as well as for use on consultation-liaison services. This volume provides a good background but is more of a well-rounded, thoughtful collection than an ideal textbook. For didactic purposes, selected chapters would do well, including Dr. Engel’s two reprinted papers.

References


WILLIAM M. GREENBERG, M.D.
Ramsey, N.J.


In the epistemological spirit of trying to make (human) order out of (natural) disorder, this ambitious book attempts to trace the logic-oriented striving of psychology and psychiatry to apply observation, statistics, and logic to the chaotic phenomena of the mentally ill mind. For Americans, the use of British spellings (“behaviour”) and grammatical constructions (“the latter became subject to much philosophical stick”) may be distracting. For believers in biological underpinnings, descriptions such as, “The person appears to have nothing to be depressed about, or can think of no reason to be depressed,” may elicit head-shaking outrage. Clinicians seeking research or even anecdotal evidence of best practices will be entirely disappointed because these pages offer abstractions only and no immediately applicable hypotheses or data. However, if the reader is open-hearted and willing to adventure into the overlapping realms of theoretical philosophy and medical nosology, then there may be surprising rewards at the conclusion. Initial familiarity with the treatises of Plato, Michel Foucault, Karl Jaspers, Wittgenstein, Hume, and Thomas Szasz will be helpful, as will a schematic overview of neurobiology.

The type of analysis in Mind, Meaning, and Mental Disorder may be demonstrated by a series of reasoning statements and paradigms. For example, if one’s former happy internal state is presumed to be linked to a happy external event, and one’s current unhappy state cannot be associated with any precedent unhappy stimulus, then could both the happiness and unhappiness be random rather than linked to or evoked by biochemistry or genetics or contagion? The conclusion of the authors is that only brain processes are genuinely causal. Alternatively stated, 1) intentionality is critical in the regulation and prediction of action, 2) intentional processes are causal, 3) intentional processes that regulate behavior are encoded in the brain, and 4) intentionality involves relativity. The process by which the authors arrive at their conclusion is through fording the dense and swirling waters of “meaning-as-culture”; the assumptions and methods of science; the conceptualization of the self, personality, and uniquely individual cognitive commitment; the distinctions between “understanding” and “explanation”; and the evolutionary connections between psy-

This book, part of the Norton Series on Interpersonal Neurobiology, is a follow-up to Affect Regulation and the Origin of the Self (1), which focused on regulation theory and its relation to affect and attachment and their impact on socioemotional development and psychopathology. The current book attempts to take the topic farther, with sections on Developmental Affective Neuroscience and Developmental Neuropsychiatry. The author integrates material from social, psychological, and biological realms into a comprehensive model to provide testable psychobiological models of human brain dysfunction. In the process, he presents research from different psychiatric theoretical perspectives, including psychodynamic, behavioral, and sociological frameworks. The ideas presented will stimulate clinicians and experimental scientists to explore this important topic. Although the book appears intended for those in clinical practice, it may appeal more to theorists and researchers as a synthesis of important concepts, theories, and other data. A companion book is Affect Regulation and the Repair of the Self (2).

Section 1, Developmental Affective Neuroscience, provides an overview for the book (e.g., regulation theory, perspectives from others on the topic). The concept of self-regulatory systems is introduced and applied to brain development. Indubitably, the development of the infant’s brain occurs in the context of a relationship with another self (e.g., the mother). Thus, a central theme is the experience-dependent relationship be-
between infant and mother for the development of self-regulatory processes. The work of Bowlby and others on attachment is integrated with neurobiological findings. In particular, the function of the right prefrontal cortex, parent-infant communication, and attachment disorders are discussed in detail. When available, functional magnetic resonance imaging findings are included. Section 2, Developmental Neuropsychiatry, discusses predispositions to psychiatric disorders. Separate chapters review how brain development, affect regulation, and health might be affected by 1) a secure attachment, 2) relational trauma, and 3) eventual affect dysregulation.

We have several suggestions for the next edition. First, we would like to see case examples to apply the information, interweave the presented theories, and “spell out” the concepts in a more relevant manner. This change would make it more readable and more directly applicable clinically. Second, the chapters could benefit from a standardized format, with an introduction, objectives, background information, and a summary of key points. This change would also make it more useful as an educational text. Third, in this complex discussion of the intersection of many fields and particularly neuroscience, it is surprising that few figures or tables are used to help the reader. Finally, particular emphasis is placed on the relation between structure and function of the right hemisphere during critical stages of development. Unfortunately, supporting neurobiological evidence is weak (e.g., regarding the specificity of abnormal development of the right prefrontal cortex), and original references are sometimes overinterpreted. This pitfall, however, surely reflects the paucity of neuroanatomical information on the normal development of the primate brain.

This is an interesting, useful integration of data from many fields—all intersecting in the areas of attachment, affect, and pathology—that makes it a very unique reference for clinicians who want to explore the theoretical underpinnings of these areas. It is an ambitious work written by a knowledgeable author. It is already good, but with the changes noted it would be outstanding in a second edition.

References


DONALD M. HILTY, M.D.
BLYTHE A. CORBETT, Ph.D.
Pierre Lavenex, Ph.D.
Sacramento, Calif.


Working in the Yale Psychology Department, Sidney Blatt has tenaciously pursued a formulation of depression for more than 30 years. This book captures his journey and charts the development of the model from personal sources (early wrestling with Freud’s “Mourning and Melancholia” [1] and analytic observations), through years of applied research, to this proposal of an integrated model.

The model argues for two types of depression but is not the classic binary (i.e., endogenous versus neurotic/reactive) one. In essence, Blatt argues for developmentally determined “anaclitic” or dependent depression (where the depressed individual’s preoccupations involve themes of abandonment and loss) and “introjective” depression (typified by punitive and harsh self-criticism), with differential developmental factors creating the shared vulnerability to depression. As acknowledged by Blatt, the model overlaps with views of other theorists, including Bowlby and Beck (the latter used concepts of “sociotropy” and “autonomy”).

The model is multilayered. The two types are held to reflect developmental disturbances in early parent-child relationships, frame the experiential world of those who develop depression, involve a stress-diathesis congruency model linking contextual stressful events with depression onset, direct treatment approaches, and allow “depression” to be modeled parsimoniously. The last issue is perhaps the most worrisome plank in Blatt’s argument. Arguing first that depressive subtyping has failed to be achieved by use of symptoms and that “depression” is a continuum ranging from “transient dysphoric” responses to “serious distortions of reality” (p. 29), Blatt then argues for his dichotomous model as providing a “major subtyping of depression” that is both reliable and has “demonstrated validity” (p. 30). Such claims are not supported by appropriately designed or definitive analyses. Reviewed analyses seek to affirm the model rather than refine it by refutability studies.

Where are the twin studies to consider the genetic and environmental contribution to such personality styles? Why two constructs as against at least four factors of the Big Five model? Is “anaclitic” not a component domain of neuroticism and, as measured, weighted to the well-researched but here essentially ignored facet of rejection sensitivity? If Blatt’s continuum model is valid, how necessary and sufficient are the constructs to distinguish psychotic from melancholic depression as well as the depressions experienced by psychology students from those of resilient control subjects? In chapter 8 (“Therapeutic Implications”), how does the model extrapolate to the world of clinical psychiatry (e.g., the perfectionistic individual with a psychotic depression)? Here it would appear that psychotherapy is viewed as necessary and sufficient.

There can be little doubt that developmental experiences and personality styles increase the risk to certain depressive disorders, shape the clinical picture to some degree, and have some differential impact on varying treatment modalities, and that Blatt’s contribution to developing such a matrix has and will continue to be distinctive. However, his exposition in this book posits the model as all-explanatory and has Procrustean overtones.

Reference


GORDON PARKER, M.D., Ph.D., D.Sc.
Sydney, Australia

http://ajp.psychiatryonline.org

Am J Psychiatry 162:7, July 2005