Contents: Volume 162, Issue 5

In This Issue:

Am J Psychiatry 2005 162: A64.

Perspectives

Editorials:

Glen O. Gabbard

Personality Disorders Come of Age

Reviews and Overviews:

John S. March, Susan G. Silva, Scott Compton, Mark Shapiro, Robert Califf, and Ranga Krishnan

The Case for Practical Clinical Trials in Psychiatry

Richard Mayou, Laurence J. Kirmayer, Greg Simon, Kurt Kroenke, and Michael Sharpe

Somatoform Disorders: Time for a New Approach in DSM-V

Images in Neuroscience:

Robert W. Greene

Hippocampus, V: Studying Hippocampal Behaviors

Introspections:

Shyam K. Bhat

How Many Psychiatrists Does It Take?

Clinical Case Conference:
Glen O. Gabbard and Tanya J. Bennett

Dilemmas in the Psychotherapy of Sexually Impulsive Patients

Images in Psychiatry:

Peter J. Buckley

William Blake (1757–1827)

New Research

Articles:

Carolyn Zittel Conklin and Drew Westen

Borderline Personality Disorder in Clinical Practice

Rocco Crino, Tim Slade, and Gavin Andrews

The Changing Prevalence and Severity of Obsessive-Compulsive Disorder Criteria From DSM-III to DSM-IV


Two-Year Prevalence and Stability of Individual DSM-IV Criteria for Schizotypal, Borderline, Avoidant, and Obsessive-Compulsive Personality Disorders: Toward a Hybrid Model of Axis II Disorders

Ephi Betan, Amy Kegley Heim, Carolyn Zittel Conklin, and Drew Westen

Countertransference Phenomena and Personality Pathology in Clinical Practice: An Empirical Investigation

Richard J. Brown, Anette Schrag, and Michael R. Trimble

Dissociation, Childhood Interpersonal Trauma, and Family Functioning in Patients With Somatization Disorder

Deborah E. Lynn, Gitta Lubke, May Yang, James T. McCracken, James J. McGough, Janeen Ishii, Sandra K. Loo, Stanley F. Nelson, and Susan L. Smalley

Temperament and Character Profiles and the Dopamine D4 Receptor Gene in ADHD
Association of the Serotonin Transporter Gene With Smoking Behavior

Thalamic and Prefrontal FDG Uptake in Never Medicated Patients With Schizophrenia

Shitij Kapur, Tamara Arenovich, Ofer Agid, Robert Zipursky, Stacy Lindborg, and Barry Jones
Evidence for Onset of Antipsychotic Effects Within the First 24 Hours of Treatment

Risperidone and Haloperidol in First-Episode Psychosis: A Long-Term Randomized Trial

David C. Henderson, Paul M. Copeland, Tara B. Daley, Christina P. Borba, Corrine Cather, Dana D. Nguyen, Pearl M. Louie, A. Eden Evins, Oliver Freudenreich, Doug Hayden, and Donald C. Goff
A Double-Blind, Placebo-Controlled Trial of Sibutramine for Olanzapine-Associated Weight Gain

Sharon Gil, Yael Caspi, Irit Zilberman Ben-Ari, Danny Koren, and Ehud Klein
Does Memory of a Traumatic Event Increase the Risk for Posttraumatic Stress Disorder in Patients With Traumatic Brain Injury? A Prospective Study

Paul E. Holtzheimer, III, Joan Russo, Douglas Zatzick, Christopher Bundy, and Peter P. Roy-Byrne
The Impact of Comorbid Posttraumatic Stress Disorder on Short-Term Clinical Outcome in Hospitalized Patients With Depression

Relief of Expressed Suicidal Intent by ECT: A Consortium for Research in ECT Study


Functional Brain Imaging Alterations in Acne Patients Treated With Isotretinoin

Brief Reports:


Higher Levels of Basal Serial CSF Cortisol in Combat Veterans With Posttraumatic Stress Disorder

Royce Lee, Thomas D. Geraci, Tri., John W. Kasckow, and Emil F. Coccaro

Childhood Trauma and Personality Disorder: Positive Correlation With Adult CSF Corticotropin-Releasing Factor Concentrations

Rachel Yehuda, Julia A. Golier, and Shira Kaufman

Circadian Rhythm of Salivary Cortisol in Holocaust Survivors With and Without PTSD

Alice Behrendt and Steffen Moritz

Posttraumatic Stress Disorder and Memory Problems After Female Genital Mutilation

Young-Gyu Chai, Dong-Yul Oh, Eun Kee Chung, Gil Sook Kim, Leen Kim, Yu-Sang Lee, and Ihn-Geun Choi

Alcohol and Aldehyde Dehydrogenase Polymorphisms in Men With Type I and Type II Alcoholism

Camila Guindalini, Sandra Scivoletto, Ricardo G.M. Ferreira, Gerome Breen, Monica Zilberman, Marco Aurélio Peluso, and Mayana Zatz

Association of Genetic Variants in Alcohol Dehydrogenase 4 With Alcohol Dependence in Brazilian Patients
Christopher Baethge, Ross J. Baldessarini, Hari-Mandir Kaur Khalsa, John Hennen, Paola Salvatore, and Mauricio Tohen

Substance Abuse in First-Episode Bipolar I Disorder: Indications for Early Intervention

Rikus Knegtering, Pepijn Baselmans, Stynke Castelein, Fokko Bosker, Richard Bruggeman, and Robert J. van den Bosch

Predominant Role of the 9-Hydroxy Metabolite of Risperidone in Elevating Blood Prolactin Levels
Am J Psychiatry 2005 162: 1010-1012.

Susan R. McGurk, Cameron Carter, Robert Goldman, Michael F. Green, Stephen R. Marder, Haiyi Xie, Nina R. Schooler, and John M. Kane

The Effects of Clozapine and Risperidone on Spatial Working Memory in Schizophrenia

Communications and Updates

Letters to the Editor:

DEBORAH R. KIM and JEFFREY P. STAAB
Quetiapine Discontinuation Syndrome

LEONA YIP, AHMED MASHHOOD, and SUNÉ NAUDÉ
Low IQ and Gasoline Huffing: The Perpetuation Cycle

MOHAMED RAHMAN and NADEEM HAIDER
Anticonvulsant Hypersensitivity Syndrome From Addition of Lamotrigine to Divalproex

DANNY SULLIVAN, MARK WALTERFANG, and DENNIS VELAKOULIS
Bipolar Disorder and Niemann-Pick Disease Type C

ALICIA E.B. LAWRYNOWICZ and TIMOTHY D. BAKER
Suicide and Latitude in Argentina: Durkheim Upside-Down
KATHARINE A. PHILLIPS
Olanzapine Augmentation of Fluoxetine in Body Dysmorphic Disorder

KIM J. MASTERS
Pilocarpine Treatment of Xerostomia Induced by Psychoactive Medications

DIRK LICHTERMANN
Cancer Risk in Parents of Patients With Schizophrenia

SUSANNE OKSBJERG DALTON, THOMAS MUNK LAURSEN, LENE MELLEMKJAER, CHRISTOFFER JOHANSEN, and PREBEN BO MORTENSEN
Dr. Dalton and Colleagues Reply

JEFFREY A. MATTES
Irritability and Depression

GIOVANNI B. CASSANO
Dr. Cassano Replies

ASHLEY SMITH, ASHOK KUMAR JAINER, and CLAIRE BENNET
Sertraline for Recurrent Major Depression

ARTHUR RIFKIN and WILLIAM DAVID RIFKIN
Psychiatry and General Practitioners

CHUAN-SHENG ZHAO, KIRSI PUURUNEN, JUHANI SIVENIUS, JUKKA JOLKKONEN, and TIMOTHY SCHALLERT
Psychotropic Medication and Stroke Outcome

MINNA M. RAIVIO, JOUKO V. LAURILA, TIMO E. STRANDBERG, REIJO S. TILVIS, and KAISU H. PITKÄLÄ
Psychotropic Medication and Stroke Outcome

NATHAN HERRMANN, KRISTA L. LANCTÔT, and MUHAMMAD MAMDANI
Dr. Herrmann and Colleagues Reply
JONATHAN FOULDS and JILL WILLIAMS
Tobacco Use and Cataracts in Patients With Schizophrenia

THOMAS E. STEELE
Tobacco Use and Cataracts in Patients With Schizophrenia

STEPHEN R. MARDER, SUSAN M. ESSOCK, ALEXANDER L. MILLER, ALAN FRIEDMAN, and MT. SINAI CONFERENCE PARTICIPANTS
Dr. Marder and Colleagues Reply

MARIO SPERANZA, MAURICE CORCOS, OLIVIER GUILBAUD, GWENOLÉ LOAS, and PHILIPPE JEAMMET
Alexithymia, Personality, and Psychopathology

HANS JOERGEN GRABE, CARSTEN SPITZER, and HARALD J. FREYBERGER
Dr. Grabe and Colleagues Reply

J. MIGUEL PENA-ANDREU, ANDRES FONTALBA, MONICA CARRERA, J. LUIS MARTINEZ, and M. ISABEL RODRÍGUEZ
Refining the Personality Disorder Diagnosis

MARK A. AMDUR
Comment on Hoarding

SANJAYA SAXENA
Dr. Saxena Replies

Book Forum:

PHILIP G. JANICAK
Practical Management of Dementia: A Multi-Professional Approach

MICHAEL R. BIEBER
Alzheimer’s Disease: The Physician’s Guide to Practical Management

DANIEL P. van KAMMEN
Schizophrenia Into Later Life: Treatment, Research, and Policy
RICHARD P. KLUFT

Brief Treatments for the Traumatized: A Project of the Green Cross Foundation

TRUCE T. ORDOÑA

Posttraumatic Stress Disorders in Children and Adolescents Handbook

DONALD D. DENTON

Posttraumatic Stress Disorder: Malady or Myth?

JUDITH L. HERMAN

Early Intervention for Trauma and Traumatic Loss

MARC GALANTER

My Name Is Bill: Bill Wilson—His Life and the Creation of Alcoholics Anonymous

MARK S. GOLD and JOHN F. BRANDT

Handbook of Clinical Alcoholism Treatment

ARON S. WOLF


WILLIAM R. FLYNN

Dual Diagnosis • Psychosocial Treatments

PAUL C. HORTON

Children of Color: Psychological Interventions With Culturally Diverse Youth, 2nd ed.

RODRIGO A. MUÑOZ

Clinician’s Guide to Cultural Psychiatry

Corrections:

Correction

Correction
Counting the Counters

The rates of obsessive-compulsive disorder (OCD) in several epidemiological studies have been surprisingly high, up to 2.5% of the general population. Broad diagnostic criteria and lay interviewers may have led to overcounting. Many people have intrusive thoughts and ritualized behaviors, but generally these do not interfere with our lives. Crino et al. (p. 876) report an Australian survey of 10,641 adults that included only obsessions or compulsions causing marked distress, consuming substantial time, and interfering with functioning. By these standards, OCD was present in 0.6% of the population. OCD was often accompanied by other psychiatric disorders, but substance abuse was less common than in other surveys. Compared to earlier diagnostic systems, the current criteria were more likely to identify people who are disabled, receive medical services, and are unemployed. This suggests a more accurate identification of people with true illness.

Temperament: Missing Link Between ADHD and the DRD4 Gene?

The quest to discover genes underlying attention deficit hyperactivity disorder (ADHD) has led researchers to the gene for the dopamine D4 receptor (DRD4), particularly the 48-base-pair (48-bp) variant. As DRD4 has also been associated with the trait of novelty seeking, Lynn et al. (p. 906) wondered whether novelty seeking is an intermediary between the DRD4 gene and ADHD. Apparently not, according to personality profiles and genetic tests of 171 adults with children with ADHD. The 7-repeat allele of the DRD4 48-bp variant was present in 33% of these parents. It predicted the presence of ADHD but not the trait of novelty seeking. However, novelty seeking contributed to ADHD even more than did DRD4. Thus, its association with ADHD is strong but appears not to stem from the 48-bp DRD4 variant.

The Smoking Gene

Susceptibility to smoking may be influenced by the gene for the serotonin transporter, a protein that controls cellular movement of serotonin. Production of the protein is affected by a promoter region on this gene and by certain DNA sequences with variable numbers of tandem repeats (VNTRs). The serotonin transporter gene may also influence smoking indirectly, by determining personality traits that influence smoking. In 330 families of smokers and nonsmokers, Kremer et al. (p. 924) found that smoking was strongly associated with the long allele for the promoter region of the serotonin transporter gene and with the 12-repeat VNTR. Novelty seeking and reward were weakly related to the genetic variations but were independent of smoking. The relationship between smoking and the serotonin transporter gene did not vary by the degree of nicotine dependence, suggesting that the gene influences whether a person begins smoking, rather than the duration or quantity of smoking.

The Drug-Free Thalamus in Schizophrenia

The thalamus comprises multiple nuclei that filter information from the senses and relay it to the prefrontal cortex and elsewhere. Patients with schizophrenia have disturbed sensory and attentional functioning, pointing to problems in the thalamus. Studies of how the thalamus works in schizophrenia have been complicated by patients’ previous antipsychotic treatment, use of mental tasks that did not stimulate the thalamus, and lower-resolution brain imaging. Leherer et al. (p. 931) used positron emission tomography to compare brain metabolism in patients awaiting treatment for schizophrenia and in healthy subjects as they performed a visual attention task that activates the thalamus. The results confirm the previous finding of lower activation in the medial and posterior thalamus of schizophrenia patients than in healthy subjects. The affected regions include the medial dorsal nucleus and the pulvinar, which have important reciprocal connections with prefrontal and temporal regions.

Some Things Are Better Not Remembered

Many people with posttraumatic stress disorder (PTSD) have frequent, disturbing memories of the traumatic event, but amnesia also can occur. To assess whether remembering, or not remembering, the event is related to the development of PTSD, Gil et al. (p. 963) questioned 120 patients with traumatic brain injuries within 24 hours of the event; 55 remembered the traumatic event and 65 did not. Six months after their injuries, 14% had PTSD, and it was twice as likely among the patients who remembered what happened as among those who did not. Patients with postinjury memories were more likely to reexperience the event at 6 months, but not to have avoidance and hyperarousal symptoms. Asking patients whether they remember the event may help identify those at risk of PTSD. On the other hand, interventions that elicit traumatic memories may be counterproductive in some cases.
Personality Disorders Come of Age

Personality disorders have often been relegated to stepchild status within psychiatry. Insurance and managed care companies may incorrectly assert that they are not treatable and, therefore, that treatment of these patients is not reimbursable. Psychiatrists themselves often confine their diagnoses to axis I syndromes. Research dollars for randomized, controlled trials of personality disorder treatments have been hard to come by.

A quarter-century after the creation of the DSM axis II, however, personality disorders have come of age. They have their own international organization devoted to studying them, and treatments of proven efficacy have been developed (1, 2). A respected personality disorders journal has been in press for nearly two decades. Intellectual ferment has never been more active in the personality disorders field.

Three articles in this issue of the Journal reflect this ferment and contribute to the ongoing dialogue about the future direction of personality disorders, especially in light of the anticipation of major changes in DSM-V. McGlashan and colleagues provide yet another significant contribution from the Collaborative Longitudinal Personality Disorders Study on the fate of four DSM-IV personality disorders: borderline, schizotypal, avoidant, and obsessive-compulsive. In 24-month blind follow-up assessments, the investigators were able to identify certain traits that were relatively fixed, whereas other criteria appeared to be more reactive and behavioral. In borderline personality disorder, for example, the authors suggested that the more stable criteria, such as anger and impulsivity, may represent the biogenetic core of borderline personality disorder, and the identification of these features may help with the modification of diagnostic criteria in future renditions of the DSM. They also speculated that the least stable criteria, such as self-injury and abandonment concerns, may be better targets for psychosocial interventions, while the core biological criteria may be the best targets for biological treatments. These suggestions have to be considered tentative, however, because, despite the authors’ efforts to take treatment into account (3), the details of which patients received which treatments were not specified.

Zittel Conklin and Westen provide another type of data about borderline personality disorder. In a continuation of previous work (4–6), these investigators sought to characterize borderline personality disorder patients in the community, compared to those who are studied in academic centers. Using the Q-sort method of providing personality descriptions, they found that borderline personality disorder patients seen in everyday practice appear to have more distress and emotional dysregulation than what is captured by the DSM-IV criteria. The two items most descriptive of the borderline personality disorder patients in their study were “tends to feel unhappy, depressed, or despondent” and “emotions tend to spiral out of control.”

These findings will resonate with psychiatrists who attempt to treat this group of patients. Often people with borderline personality disorder are dismissed as “manipulators” or regarded pejoratively as “splitters.” What these findings underscore is that these people are in pain and feel that they are at the mercy of a maelstrom at the core of their being. Clinicians must be trained to recognize this pain and to get beyond the negative and alienating features of borderline personality disorder patients in order to endure the emotional roller coaster ride that often accompanies the treatment.

“The investigators emphasize the value of data provided by experienced clinical observers who see a patient over time.”
The investigators emphasize the value of data provided by experienced clinical observers who see a patient over time. Research instruments that assess an individual at one snapshot in time are fraught with problems in the assessment of personality disorders (7). Patients with borderline personality disorder may be kaleidoscopically different from one week to the next based on their affective state and the vicissitudes of their object relationships (8).

A classic *New Yorker* cartoon from the early 1960s depicts a peacock with its spectacular tail in full splendor saying to a smaller bird with no tail whatsoever, “Now let’s talk about you.” The humor in the cartoon derives from the fact that every reader knows how it feels to be on the receiving end of a narcissistic display of self-importance. Indeed, in the third contribution on personality disorders featured in this issue of the *Journal*, Betan et al. report an empirically based description of countertransference responses to narcissistic patients that strongly resembles theoretical and clinical accounts. In a random sample of 181 psychiatrists and clinical psychologists from North America, the investigators tested a new questionnaire and found that it yielded eight clinically and conceptually coherent factors that were independent of the clinicians’ theoretical orientation. As one might anticipate, the eight factors were associated in predictable ways with axis II pathology. As part of their data analysis, they created a composite description of countertransference patterns in the treatment of patients who met the criteria for narcissistic personality disorder. They found that clinicians reported feeling resentment, anger, and dread in their interactions with the patient and tended to feel devalued and criticized by the patient. During their appointments with such patients, they felt distracted and avoidant and wished to end the treatment.

Clinicians have long known that patients with personality disorders re-create their characteristic mode of relatedness in their relationship with the clinician and impose a certain way of thinking, feeling, and reacting on the clinician. A problem for clinicians in systematically using this information diagnostically is that countertransference draws from the clinician’s own conflicts and past experiences as well as from the feelings induced by the patient (9, 10). Nevertheless, what the data from this investigation illustrate is that there is an “average expectable countertransference” that may transcend the highly specific individual feelings brought to the clinical setting based on one’s own personal background. Professionals who work in group treatment settings, such as day treatment units or day hospitals, know that there are consistent reactions to certain types of patients, reflecting potential problems in the treatment.

A dilemma posed by the contemplation of including countertransference responses as an aid to the diagnosis of personality disorders, however, is that some forms of countertransference are largely unconscious. Often clinicians become aware of their feelings toward the patient only through small enactments, such as starting appointments late, getting sleepy to the point where the patient notices it, or making sarcastic comments in the guise of confrontation. Hence countertransference may be a *discovery* based on careful self-scrutiny that emerges in the course of the treatment.

Finally, if, as Betan et al. suggest, countertransference should be given the position of importance that it so richly deserves in the understanding and treatment of psychiatric patients, a formidable obstacle must be addressed. Self-reflection is no longer emphasized in residency training programs as it was in decades past. Trainees are not necessarily encouraged to have a personal psychotherapy experience to examine their own conflicts. Hence the realm of countertransference may be an unexplored continent. Psychiatry has long distinguished itself from other medical specialties by its attention to the clinician’s feelings as an important diagnostic and therapeutic tool in its armamentarium. To effectively treat patients with personality disorders, that tool must not disappear through disuse atrophy.
References


Glen O. Gabbard, M.D.

Address correspondence and reprint requests to Dr. Gabbard, Baylor College of Medicine, 6655 Travis, Suite 500, Houston, TX 77030; ggabbard12@aol.com (e-mail).
It is widely acknowledged that the cultural shift toward evidence-based practice that is now under way in psychiatry (1, 2) requires that research be seen as integral rather than as separate from clinical practice (3). It is less widely acknowledged that using high-quality research evidence to guide clinical practice has as many implications for researchers as it does for clinicians (4, 5). Although substantial progress has been made in developing and testing new treatments, researchers for the most part have failed to design trials that maximize clinical utility for practicing clinicians and other decision makers, or, stated in terms of the contrary position, current best evidence often is not perceived by decision makers as relevant to clinical practice, thereby substantially diluting its impact (6). Moreover, clinical research in psychiatry is hobbled by high costs, lack of funding, regulatory burdens, fragmented infrastructure, slow results, and a shortage of qualified investigators and willing participants. To speed the translation of clinical research into medical practice, the Director of the National Institutes of Health (NIH), Elias Zerhouni, M.D., recently emphasized the need for practical clinical trials that are explicitly designed to aid decision makers who are faced with choices about patient care, whether at the doctor-patient level or the policy level (7). Accordingly, we make the case in this article that adopting the practical clinical trials model in psychiatry would have a positive public health impact in the care of mentally ill children, adolescents, and adults.

What Defines a Practical Clinical Trial?

More than a decade ago, Sir Richard Peto and colleagues coined the term “large, simple trial” (8) and, in so doing, emphasized that treatment outcomes studies should have sufficient power to identify modest clinically relevant effects, should employ randomization to protect against bias, and should be simple enough to make participation by patients and providers reasonable (9). More recently, Tunis et al. (5) described the defining features of a practical clinical trial; they are a study design that compares clinically important interventions, a diverse population of study participants representative of clinical practice, inclusion of a range of heterogeneous practice settings that are also representative of clinical practice, and measurement of a broad range of clinically relevant health outcomes. (Many consider the terms “practical clinical trial,” “pragmatic trial,” and “large, simple trial” to be interchangeable, because all three terminological conventions designate research that shares the primary goal of asking a question that, when answered, will inform persons who are making decisions about the care of patients. We use the appellation “practical clinical trial” because it
is the emerging standard term for trials of this type.) Starting with the definition offered by Tunis et al. (5) and broadening it to include principles articulated by Peto and Baigent (9) and others (4, 10, 11), practical clinical trials can be characterized by eight defining principles:

- **Questions must be simple, clinically relevant, and of substantial public health importance.** Each practical clinical trial begins with a good question (12), for example, “Does a widely practicable treatment (or treatments) alter the major outcome(s) for a common mental disorder?” Not coincidentally, the more common and important the question, the easier it is to enroll subjects in a practical clinical trial.

- **Practical clinical trials are performed in clinical practice settings.** To enhance generalizability, a clinically representative (typically heterogeneous) sample of patients is recruited and treated in the clinical practice settings in which these patients are ordinarily found (13).

- **Study power is sufficient to identify small to moderate effects.** Although the size of a practical clinical trial depends on the question that needs to be answered, practical clinical trials are usually larger (and almost always simpler) relative to efficacy trials. For example, to identify clinically relevant if modest effects in a head-to-head trial of two active treatments, a practical clinical trial might allocate thousands of patients in hundreds of clinical centers to different treatments just as they are delivered in community settings (13).

- **Randomization is the best defense against bias.** Experimental manipulation of treatment assignment (e.g., randomization) is the best defense available against selection bias and confounding factors (14). Thus, in practical clinical trials, patients are randomly allocated to one or more active treatments and, depending on the question, to a control or active comparison condition or both.

- **Randomization depends on clinical uncertainty.** A good question is interesting only if it represents an important area of clinical uncertainty (2). As applied to a practical clinical trial, the uncertainty principle means that there is no substantial empirical basis for preferring one or another possible treatment (including study and nonstudy treatments) for a particular patient (9, 13). In turn, true uncertainty regarding what is the best treatment makes participation in research ethically reasonable with respect to the balance between benefit and harm and limits the possibility that real or apparent bias regarding participation will arise from the doctor-patient relationship.

- **Outcomes are simple and clinically relevant.** Outcomes are tracked by using unambiguous, readily detectable endpoints that reduce classification errors and simplify data collection (15, 16). A simple way of considering whether an outcome is appropriate for a practical clinical trial is by determining whether it 1) can easily be recognized by clinicians and/or patients and 2) reflects the construct of living longer, feeling better, avoiding unpleasant experiences such as having to go to the hospital, and spending less money.

- **Assessments and treatments enact best clinical practice.** To encourage findings that are clinically meaningful and that will transfer readily to clinical practice, practical clinical trials are characterized by simple, straightforward assessment and treatment methods typical of everyday good clinical practice (17).

- **Subject and investigator burden associated with research goals are minimized.** Although the degree of simplicity is dependent on the aims of the trial, practical clinical trials are completed in a timely fashion at relatively low cost by employing simple methods of protocol development, protocol dissemination, data gathering, and quality assurance (15, 16). In particular, the number and type of data elements (and, hence, subject and investigator burden) are kept small and straightforward so as not to discourage provider or patient participation and to maximize the number of subjects per dollar spent.

**Scope of Practical Clinical Trials**

Granting that the essential components of an outcomes study necessarily will vary as a function of the question being asked (18), practical clinical trials may be distinguished from efficacy trials, hybrid efficacy/effectiveness trials, dissemination trials, and practice research (Table 1).

In mental health outcomes research, the distinction is frequently made between efficacy and effectiveness trials (6), although Kraemer (19) cogently argued that efficacy and effectiveness represent anchor points at the end of a continuum rather than discrete categories. Randomized, controlled trials that focus on therapeutic efficacy ask the question: “Will a treatment work under ideal conditions?” With the industry-funded registration trial as the paradigmatic example, efficacy trials employ relatively homogeneous samples and intensive assessment strategies restricted to one or two primary outcomes. In so doing, efficacy trials explicitly maximize assay sensitivity for detecting a treatment signal on the primary outcome measure with the minimum sample size in the comparison of an active treatment with a control condition. Some efficacy trials also include an explanatory component, such as a pharmacokinetic substudy, designed to elucidate how or why a treatment works; usually, such substudies are accomplished at the cost of a much higher density of assessment. In contrast, effectiveness trials focus on gathering information of maximum interest to clinicians and other decision makers. In so doing, effectiveness studies, broadly conceptualized, employ heterogeneous samples that are recruited in a variety of practice settings and are assessed not only for primary outcomes but for a wide range of outcomes relevant to public health, such as co-
TABLE 1. Design Characteristics of Practical Clinical Trials, Compared With Efficacy Trials, Hybrid Efficacy/Effectiveness Trials, Dissemination Trials, and Practice Research

<table>
<thead>
<tr>
<th>Study Characteristic</th>
<th>Practical Clinical Trial</th>
<th>Efficacy Trial</th>
<th>Hybrid Efficacy/Effectiveness Trial</th>
<th>Dissemination Trial</th>
<th>Practice Research</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Research question</strong></td>
<td>Will a treatment do more good than harm under best-practice clinical conditions?</td>
<td>Will the treatment work under ideal conditions?</td>
<td>Will treatment work in a heterogeneous sample under research sampling and assessment conditions?</td>
<td>Which method is more effective in disseminating a proven treatment to practice?</td>
<td>Which treatments are being used with what outcomes?</td>
</tr>
<tr>
<td><strong>Setting</strong></td>
<td>Clinical practice</td>
<td>Specialized center</td>
<td>Specialized academic and clinical centers</td>
<td>Typically, large group practice settings</td>
<td>Clinical practice</td>
</tr>
<tr>
<td><strong>Design</strong></td>
<td>Randomized clinical trial</td>
<td>Randomized clinical trial</td>
<td>Randomized clinical trial</td>
<td>Randomized clinical trial</td>
<td>Observational</td>
</tr>
<tr>
<td><strong>Sampling frame</strong></td>
<td>Usual patients, subject to uncertainty principle</td>
<td>Highly selective</td>
<td>Less selective</td>
<td>Less selective</td>
<td>Practice settings</td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td>Depends on the question, but typically ≥1,000 subjects per condition</td>
<td>Typically &lt;200 subjects per condition</td>
<td>Typically &lt;200 subjects per condition</td>
<td>Cluster randomized designs enrolling hundreds of patients</td>
<td>Hundreds of practices</td>
</tr>
<tr>
<td><strong>Power to find difference in response rates between two arms</strong></td>
<td>2.5%–10% difference</td>
<td>20%–30% difference</td>
<td>20% difference</td>
<td>10%–20% difference</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>Other interventions</strong></td>
<td>Prohibited, limited, or local practice</td>
<td>Prohibited</td>
<td>Prohibited</td>
<td>Prohibited or limited</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>Control</strong></td>
<td>Placebo, treatment as usual, or active comparator</td>
<td>Typically placebo</td>
<td>Placebo or treatment as usual</td>
<td>Treatment as usual</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>Assessments</strong></td>
<td>Simple and clinician-friendly</td>
<td>Elaborate</td>
<td>Elaborate</td>
<td>Simple and clinician-friendly</td>
<td>Survey research methods</td>
</tr>
<tr>
<td><strong>Patient subgrouping (moderators/predictors of outcome)</strong></td>
<td>Power acceptable for moderators, enhanced generalizability</td>
<td>Limited by low power</td>
<td>Primarily hypothesis-generating</td>
<td>Limited by low power</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>Explanatory/mediational mechanisms</strong></td>
<td>Limited to compliance</td>
<td>Data-intensive assessments</td>
<td>Limited by low power and inadequate measurement density</td>
<td>Depends on trial aims</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>Data collection and monitoring</strong></td>
<td>Simple</td>
<td>Elaborate</td>
<td>Elaborate</td>
<td>Simple</td>
<td>Summary questionnaires, chart review</td>
</tr>
<tr>
<td><strong>Quality assurance</strong></td>
<td>Minimal</td>
<td>Elaborate</td>
<td>Elaborate</td>
<td>Intermediate</td>
<td>Intermediate</td>
</tr>
<tr>
<td><strong>Study management</strong></td>
<td>Minimal</td>
<td>Intensive</td>
<td>Intensive</td>
<td>Intermediate</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>Generalizability</strong></td>
<td>High</td>
<td>Low</td>
<td>Intermediate</td>
<td>High</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

morbidity, quality of life, and cost-effectiveness (20, 21). The current generation of large comparative treatment trials funded by the National Institute of Mental Health (NIMH) (see, for example, reference 22) typically combines efficacy (usually involving intensive assessments) and effectiveness (usually a broad sampling frame) elements and thus occupies an intermediate position on the efficacy-to-effectiveness continuum.

Located at the far end of the effectiveness part of the spectrum, practical clinical trials ask, “Will a treatment do more good than harm under usual conditions of clinical practice?” Consequently, practical clinical trials are not well suited to outcomes studies with efficacy and explanatory aims, including treatment development studies (especially for novel treatments), dose-finding studies, studies of mechanisms of treatment action involving large numbers of mediator variables, or acute pharmacokinetic studies, some but not all of which necessarily precede practical clinical trials. Practical clinical trials are best suited to tests of the effectiveness of newly introduced treatments, compared to standard treatment; assessment of treatment additions to address partial response; research designs with many safety outcomes; population pharmacokinetic studies; and studies of events with relatively rare but important outcomes, such as death or rehospitalization.

In addition, because of enhanced sample representativeness and large sample sizes, which provide enhanced power to identify small effects, practical clinical trials are the ideal venue to investigate the question, “Which treatment for which patient with what subgrouping characteristics?” (5, 23). For example, practical clinical trials, unlike studies with smaller and less heterogeneous samples, typically have ample power to examine subgrouping variables, such as race, gender, age, severity, and pattern of comorbidity, alone and in combination. (One practical clinical trial of coronary reperfusion strategies spawned more than 80 papers, most of them involving subgrouping...
analyses [24].) Likewise, practical clinical trials that include genetic predictor variables that have low base rates will be especially valuable as pharmacogenetics assumes a more prominent role in designing clinical trials and in guiding the selection of treatments (25).

It is important to note that practical clinical trials as applied to safety outcomes constitute a major means to address the balance between benefit and harm in treatments as they are administered in clinical practice (26). The prevalence and relative risk of common adverse events can be estimated far more precisely in a 2,000-subject trial than in a 200-subject trial. Identifying the relative risk of uncommon (1 in 500 to 1,000) or rare (1 in 10,000) adverse events is possible only in the practical clinical trial setting. In this regard, adverse event reporting in practical clinical trials has been shown to enhance safety during the trial and to facilitate the role of data monitoring committees and institutional review boards confronted with multiple adverse event reports (27, 28).

Unlike dissemination research, which focuses on assessing the effectiveness of different dissemination strategies, practical clinical trials focus on comparing treatments rather than on the dissemination of those treatments. However, a key barrier to disseminating evidence-based treatments is the lack of development and testing of treatments in real-world practice settings (29). In this context, a preexisting practical clinical trial can facilitate dissemination of a new treatment through the development of new knowledge and of procedures that should prove readily transferable to clinical practice.

Lastly, although they take place in clinical practice settings, practical clinical trials are not the same as practice research, which is more powerful for profiling practice patterns than for asking and answering questions regarding the effectiveness of treatment (30, 31). On the other hand, it is undeniable that in fields that do practical clinical trials, the trials themselves change practice (15, 16). Those who participate in trials are more likely to practice according to the results of those trials, and the fact that a trial can be done in a typical practice situation greatly enhances the ability to disseminate the technology. Thus, practice research involving a practical clinical trial network typically focuses on understanding the practice settings to which the results of a practical clinical trial will generalize and the effects of participation in practical clinical trials on practitioners in those settings (5).

Examples of Practical Clinical Trials

**Practical Clinical Trials in Medicine**

Over the past several decades, the pressing need to understand the effect of widely practiced treatments on the major outcomes for common illnesses such as heart attack and stroke gave rise to increasingly large and sophisticated practical clinical trial networks in cardiovascular medicine (9). This process was greatly facilitated by network development and maintenance activities funded by NIH (32). As a result, practical clinical trials are now the norm in areas such as acute coronary syndrome and heart failure (10). For example, the Virtual Coordinating Center for Global Cardiovascular Research network has so far conducted 15 trials of thrombolytic agents with more than 150,000 enrolled patients at 3,000 sites in 50 countries (33). These practice-changing practical clinical trials have led to convincing improvements in coronary reperfusion and, more important, survival in patients with acute myocardial infarction (24).

In pediatric oncology, practical clinical trials conducted in the Children's Oncology Group network have revolutionized the care of children with cancer (34). Illustrating the fact that practical clinical trials can and do take place in specialty clinics if that is where the patients are located, more than 95% of children with cancer in the United States are treated at the approximately 500 participating centers in the Children's Oncology Group network, which is funded by the National Cancer Institute. The Children's Oncology Group provides a national network of communication for researchers, care providers, and families of pediatric patients with malignant disease and conducts laboratory investigations and clinical trials of new treatments for cancer in infants, children, adolescents, and young adults. Cases seen in the network constitute, in effect, a national registry of nearly all childhood cancers in the United States. Because of the network's broad coverage, more than half of American children with cancer are entered into at least one randomized Children's Oncology Group trial, which in turn has led to dramatic improvement in pediatric cancer outcomes.

Practical clinical trials sometimes reveal that widely used, seemingly sensible clinical strategies that have been shown to work in small industry-funded registration trials can, in actual fact, be harmful when applied in real clinical settings. In a sobering and heuristically valuable example, the Cardiac Arrhythmia Suppression Trial examined Class I antiarrhythmic drugs that had been shown in small efficacy trials to suppress ventricular extrasystoles in patients at increased risk of death after a heart attack (35). Intuitively, it seemed reasonable that a reduction in irregular heartbeats also should result in reduced mortality. However, the Cardiac Arrhythmia Suppression Trial revealed that these drugs were associated with a 250% increase in short-term mortality, whereas newer Class III antiarrhythmic agents appeared to reduce mortality (36). The Cardiac Arrhythmia Suppression Trial nicely illustrates how practical clinical trials can be useful in confirming that the results of small efficacy trials will translate (or will not translate, as the case may be) to real-world practice (37), especially when the initial trials involved surrogate markers for relatively rare outcomes such as mortality.
Practical Clinical Trials in Psychiatry

Practical clinical trials are just beginning to occur in psychiatry (1, 38). Projects illustrating recent progress in implementing practical clinical trials in the mental health field include the British Bipolar Affective Disorder: Lithium Anticonvulstant Evaluation trial involving adults with bipolar disorder (2); the industry-funded International Suicide Prevention Trial in schizophrenia (39); the NIMH-funded practical clinical trials in adults with depression, bipolar disorder, psychosis, and Alzheimer's disease (40-43); and the NIMH-funded Child and Adolescent Psychiatry Trials Network (38).

Funded by the Stanley Foundation and the British Medical Research Council, the Bipolar Affective Disorder: Lithium Anticonvulstant Evaluation trial addresses the question, “In adult patients with bipolar illness, is lithium, sodium valproate, or a combination of the two more effective in reducing the risk of further episodes of mania or depression and in improving long-term effects on mood and more tolerable in terms of side effects?” The trial will recruit thousands of patients at 54 centers who will be randomly assigned to best-practice active treatments and assessed with simple, clinically relevant categorical and dimensional outcomes measures (2).

The recently completed International Suicide Prevention Trial compared the effects of clozapine and olanzapine in 980 adults with schizophrenia at high risk for suicide at 56 sites in 11 countries (44). As is typical of a practical clinical trial, the study included freedom to augment these treatments as needed, blind ratings in which suicide was the primary outcome, and equivalent clinical contact for the two treatment groups. Clozapine proved superior to olanzapine in reducing suicidality. Although it is too soon to evaluate the impact of the International Suicide Prevention Trial on clinical practice, the investigators suggested that clozapine should be considered for all patients with schizophrenia at high risk for suicide (44).

NIMH recently funded three large practical clinical trials designed to address questions of major public health relevance that emerged from but were not answered by earlier industry-funded efficacy trials: 1) Sequenced Treatment Alternatives to Relieve Depression, which compares active treatments and treatment strategies in adults with depression (40); 2) Systematic Treatment Enhancement Program for Bipolar Disorder, which focuses on the pharmacological and psychosocial treatment of bipolar illness (41); and 3) Clinical Antipsychotic Trials of Intervention Effectiveness, which contrasts various atypical antipsychotics as treatments for psychosis or for Alzheimer's disease (42, 43). These trials have innovative experimental designs—including an equipoise-stratified design in Sequenced Treatment Alternatives to Relieve Depression (45) and an outcomes-based rerandomization scheme in Clinical Antipsychotic Trials of Intervention Effectiveness (46)—and are much larger than any study previously attempted in psychiatry. They pioneer the construction of the kind of networks that are essential for making practical clinical trials a routine part of psychiatric practice.

A partnership between researchers at Duke University and the American Academy of Child and Adolescent Psychiatry, the NIMH-funded Child and Adolescent Psychiatry Trials Network is an investigator-initiated proof-of-concept effort to establish a practical clinical trial network in pediatric psychiatry (38). Guided in the selection of specific questions by network members and expert advisory panels, the Child and Adolescent Psychiatry Trials Network will focus on two critically important clinical issues: 1) obtaining evidence from randomized trials of the effectiveness of widely used but understudied combined drug treatments and 2) investigating the short- and long-term safety of pharmacotherapy. With 200–400 child and adolescent psychiatrists each participating in two or three practical clinical trials at any one time, the Child and Adolescent Psychiatry Trials Network promises to advance both the evidence base and the research capacity of child and adolescent psychiatry.

Challenges in Adopting Practical Clinical Trials in Psychiatry

Network Construction and Maintenance

Practical clinical trials typically flourish in networks that are developed and maintained by research centers in productive partnership with clinical and administrative decision makers in federal, community, and academic settings (5, 37, 47). No stably funded practical clinical trial infrastructure currently exists in psychiatry—the current practical clinical trials in psychiatry (Sequenced Treatment Alternatives to Relieve Depression, Systematic Treatment Enhancement Program for Bipolar Disorder, Clinical Antipsychotic Trials of Intervention Effectiveness, and Child and Adolescent Psychiatry Trials Network) will sunset if their funding is not renewed. Thus, the development of a stable practical clinical trial infrastructure, including establishing a coordinating center or centers, remains a necessary step toward the implementation of the practical clinical trials model in mental health care. A coordinating infrastructure with access to researchers and other individuals with the skills and equipment required to conduct a practical clinical trial is necessary 1) to recruit and sustain a broad-based network of clinical sites representative of current practice and service systems (site development); 2) to institute the mix of financial, practical (e.g., continuing medical education), and altruistic incentives that makes ongoing network participation reasonable and desirable; 3) to develop trial-independent and trial-dependent clinical research training for site investigators; and 4) to characterize the practice patterns and factors that affect clinical decision making, treatment provision, and conformance with evidence-based treatment recommendations within each site. It is important to note that to cover the practice landscape in which mentally ill patients...
are found and to model rational triage strategies, it will be necessary to extend practical clinical trial networks to specialty clinics (which, by analogy to specialty cancer clinics, might be appropriate for rare conditions such as schizophrenia or complex tic disorders) and to community practice settings (for more common illnesses, such as attention deficit hyperactivity disorder or major depression). Other important aspects of network coordination include the development of systems for data management and the provision of analytic and dissemination infrastructures. Finally, initiatives to conduct practical clinical trials in other areas of medicine have been successful because network participants were trained in network protocols during residency and continued to participate after they began clinical practice. Hence, seamless integration of practical clinical trials into training programs (as will be done in the Child and Adolescent Psychiatry Trials Network [38]) is of paramount importance if practical clinical trials are to be widely implemented in mental health delivery systems.

**Developing Questions That Are Directly Relevant to Clinical Practice**

Although much progress has been made, clinical practice in psychiatry remains far from being evidence based (48), with heterogeneity in practice predicting heterogeneity in quality of care (49, 50). In treatment of children, for example, there has been a rapid shift from a single-drug standard toward complex multidrug treatment regimens (51) for which scientific support is completely lacking (52). These gaps in the research literature, which can be identified by examining postmarketing pharmacoepidemiologic studies (53) and findings from practice research (30, 31) and by perusing expert-driven treatment guidelines (17, 54), contribute both to the rationale and, by guiding the choice of questions, to the method for prioritizing a practical clinical trial research agenda. In addition, because physicians participating in practical clinical trials are themselves interested in securing straightforward answers to a large number of highly relevant clinical questions, vetting questions and protocols through the network members is an effective way to prioritize practical clinical trials and also serves a network maintenance function by enhancing recruitment of patients into clinical trials.

**Using Placebo Controls, Active Comparators, and Blinding**

As with efficacy studies, the choice of an experimental design in a practical clinical trial depends on the theoretical and practical aims of the trial (55). When the primary purpose of a trial is to secure an unbiased evaluation of whether a treatment signal is present, an inactive control condition—typically pill placebo in pharmacotherapy trials—maximizes the so-called assay sensitivity of the experiment. Two settings in which a placebo condition would be considered for a practical clinical trial are: 1) if one or more treatments have never been shown to be superior to an inactive control treatment or 2) if it is unreasonable to infer that the results of prior randomized clinical trials of an active treatment, compared with placebo, will be replicated if tested in the same patient population and if a true tie is unacceptable for this setting. Although the optimal time for a placebo-controlled trial is early in the treatment development cycle, before widespread belief in the clinical effectiveness of a treatment accrues, a great deal of the evidence clinicians need to practice effectively involves comparisons of active treatments and treatment addition strategies targeting partial response. This requirement makes placebo-controlled practical clinical trials appropriate long after early-phase efficacy studies have been completed.

Despite the fact that some investigators believe that a null result in a trial with two active treatments cannot be interpreted in the absence of an inactive control condition, practical clinical trials without a concurrent placebo arm are relatively common. For example, if both active treatments have shown unequivocal superiority against placebo, then a practical clinical trial without a concurrent placebo control condition is generally thought reasonable, especially if use of placebo can be said to violate the principle of therapeutic uncertainty (56). Even when this standard is not met, some advocates of practical clinical trials would argue that, unless a newer treatment has substantial advantages with respect to adverse events, user friendliness, or cost, if it cannot beat a proven treatment in a practical clinical trial powered for a small effect size, “Who needs it?”

Finally, in a practical clinical trial with an active comparator or comparators but no placebo group, the investigator must decide whether masking (blinding treatments to eliminate observer bias) is essential to the study aims. Observer bias occurs when information held by the observer influences the way he or she interprets and scores a primary outcome. In clinical practice, knowledge of treatment assignment is part of the ecological validity of the outcome (because patients know the nature of their treatment assignment) and, moreover, is consistent with the uncertainty principle for the study treatments. Thus, openly assigning patients to treatment does not necessarily detract from (and, in fact, may enhance) the robustness of a practical clinical trial.

**Using Interventions That Match Best Clinical Practice**

As in efficacy trials, treatment protocols in a practical clinical trial must specify the treatment(s) to be used; titration of medications for subjects through the use of fixed, flexible, or a combination of fixed and flexible dosing methods; range of starting doses; administration schedule; selection of brand-name or generic drugs; choice of rating forms and symptoms rated during titration and how this information will be used; adverse event monitoring.
and its effect on dosing strategy; and procedures for managing emergencies. In this regard, tensions exist between adherence (how well the standardized treatment protocol is followed) and competence (how well the treatment is done clinically); between tailoring treatment to the needs of the patient and using uniform, manualized, explicitly specified intervention(s); and between the desire for site flexibility and efficiency and the need for a common protocol that minimizes the potential for between-site drift in protocol implementation. To facilitate protocol acceptance and compliance, practical clinical trials evaluate these tensions through the perspective of the practicing clinician, especially when a standard of care is widely acknowledged. Thus, practical clinical trial researchers must work collaboratively with network clinicians to develop short, user-friendly protocols and manuals that are empirically grounded and clinically meaningful in order to maximize the generalizability of best practices to the community of psychiatrists caring for patients.

**Developing Simple, Ecologically Valid Diagnostic and Endpoint Assessments**

Although well-justified exceptions exist (see reference 57 for an explanatory aim linked to an important functional outcome), practical clinical trials generally eschew complex and time-consuming assessments designed to satisfy explanatory aims in favor of assessment strategies that are intended to be widely applicable to real-world practice settings (4). Specifically, practical clinical trials use unambiguous, readily detectable, ecologically valid diagnostic strategies and endpoints that reduce misclassification errors and simplify data collection, thereby easing the burden of participation on patients and clinicians (Table 2). For example, a practical clinical trial end-of-treatment assessment might use a DSM-IV diagnostic checklist, one or two short self-report scales, a checklist of functional outcomes, and summary clinician-rated measures, such as a Clinical Global Impression score, that together might take all of 30 minutes to complete in the context of a routine office visit. For practical clinical trials, the central question is whether these simplified assessments, which by definition should transfer readily to clinical practice, yield reliable and valid inferences regarding the outcome of treatment.

The most vexing objection aimed at the practical clinical trials model as applied to psychiatry is whether endpoints can be found that match this ideal (1, 4, 58). Practical clinical trials in psychiatry have been criticized because, unlike cancer and cardiovascular randomized clinical trials, which use “hard,” dramatic, and unmistakable events (such as mortality) as endpoints, psychiatry trials rely on “soft” behavioral/symptomatic outcomes (4). Although it is true that psychiatric outcomes are subject to greater measurement error, relative to, for example, mortality as an outcome, psychiatric research does have some readily available hard endpoints, such as predefined study rescue procedures, suicide attempts, treatment switching, hospitalization, school failure or truancy, job loss, or even dropping out of the trial itself. Furthermore, it is possible to overestimate the “hardness” of both medical and psychiatric endpoints—consider, for example, the misclassification of myocardial infarction in cardiology trials (59) or the difficulty of suicide adjudication in the International Suicide Prevention Trial (44)—and to underestimate the proven utility of “softer” endpoints, such as quality of life (60), global outcomes, or scores on psychometrically robust rating scales (see, for example, references 61, 62).

Even granting that measures used in a practical clinical trial are reliable and valid, critics argue that site differences and clinician-to-clinician variability in diagnostic and endpoint assessments could conceivably attenuate the ability to identify between-group differences (63). Fortunately, one of the strengths of the practical clinical trials model is the small standard errors associated with the large sample sizes (2). Put differently, the estimate of the true population mean almost certainly will be more precise in a 2,000-subject parallel-group practical clinical trial than in a meta-analysis of ten underpowered 200-subject efficacy trials. This feature is one of the reasons why most designers of practical clinical trials would prefer to collect a clinical global improvement score in hundreds or thousands of patients rather than to measure 100 variables in a small group of patients. Likewise, the ecological validity, and thus generalizability, of the result as it would translate to clinical practice depends on reproducing the variability present in best clinical practice. In this context, a large study “N,” coupled with assessment strategies that capture the normal outcome of treatment, can be seen as an essential ingredient, not a weakness, of the practical clinical trial approach.

**Using Quality Assurance Procedures That Reflect Good Clinical Practice**

The generalizability of a practical clinical trial is determined in large part by the extent to which the trial patients and the practice settings in which they are found resemble the patients and practice settings to which the results of the study are intended to apply. Filters—whether related to inclusion or exclusion criteria or to some aspect of subject or investigator burden associated with quality assurance procedures—attenuate this relationship by selectively limiting study enrollment or encouraging dropout.
Complex quality assurance procedures, such as intensive face-to-face training procedures, site visits, and real-time reliability checks, are impossible in a practical clinical trial. Instead, quality assurance in a practical clinical trial typically focuses on consistent implementation and documentation of compliance with assessment and treatment protocols by using adherence checklists. In this feature, practical clinical trials follow the approach of Peto and Baigent (9) with respect to the implementation of quality assurance:

Requirements for large amounts of defensive documentation imposed on trials by well intentioned guidelines on good clinical practice (or good research practice) or excessive audits may, paradoxically, substantially reduce the reliability with which therapeutic questions are answered, if their indirect effect is to make randomized trials smaller or even to prevent them starting.

Developing Innovative Methods

For the simple reason that practical clinical trials are new to psychiatry, it will be necessary to develop novel methods of sample recruitment and retention, new measures to broaden assessment of the effect of interventions at the individual and system levels, innovative research designs, innovative data entry and database management techniques, and new statistical analytic methods. Hence, practical clinical trials in psychiatry inevitably will encourage collaboration of methodologists from diverse academic backgrounds, including epidemiology, statistics, behavioral and social science, engineering, computer science, and public policy. Because longitudinal data are common in psychiatric research, particular attention must be paid to the application of long-time-frame random regression, pattern mixture, and propensity scoring techniques for handling missing data and other sources of loss of randomization; survival analysis models; power and sample size considerations; and concerns regarding multiple comparisons and multiple outcomes across many publications. In turn, attendant methodological innovation will increase the relevance of research findings for community stakeholders such as payers and public policy makers, as well as drive changes in methods for efficacy and hybrid efficacy/effectiveness trials.

Who Should Fund Practical Clinical Trials?

Practical clinical trials in psychiatry are in short supply partly because the primary funding sources for clinical research—NIMH, other NIH institutes addressing mental illness, and the pharmaceutical industry—have not until very recently provided financial support for individual practical clinical trials or a practical clinical trial network infrastructure. Clearly, an expanded role for practical clinical trials within the mental health care delivery system will depend on a substantial increase in public and private funding. For this increase to occur, clinical and health policy decision makers will need to make the funding of practical clinical trials a priority, including funding for infrastructure development, for studies as distinct from the infrastructure, for mechanisms of priority setting, and for methodological innovation. Although such a step is consistent with the 1999 recommendations of the National Advisory Mental Health Council (64), until recently no well-established funding mechanism for network studies has existed in the NIMH portfolio. There is an urgent need to further develop NIH/NIMH funding mechanisms specific to supporting networks of independent, skilled clinicians and researchers who are interested in sponsoring and performing practical clinical trials, especially when the potential profits are low but the scientific interest is high. In this context, it remains to be seen to what extent the recently published NIH “roadmap” (7) will spur a re-engineered psychiatric research enterprise that emphasizes practical clinical trials conducted through a partnership between NIMH, academic institutions, consumers, and community-based mental health care systems. If the transition to practical clinical trials is successful, the return on investment should be high. In comparison to the Treatment for Adolescents With Depression Study (22), which with 439 subjects will take 8 years and $17 million to complete, it is estimated that the Child and Adolescent Psychiatry Trials Network, with 400 child and adolescent psychiatrists each monitoring four to six subjects, will be able to complete a 1,600-subject trial from question to first manuscript in less than 2 years at a per trial cost (exclusive of the network’s infrastructure) of less than $2 million.

In most practical clinical trials, the cost of research but not the cost of patient care is borne by the network. This fact raises both practical and legal issues, because there is controversy in the managed care environment about whether insurance plans will even allow, much less cover, the cost of care for patients enrolled in practical clinical trials. The NIH position is clear. As outlined in the Code of Federal Regulations section that defines NIH policy on the determination and reimbursement of research patient care costs under grants (45 CFR 74, Appendix E),

The patient and/or third-party insurance usually will provide for reimbursement of charges for “usual patient care” as opposed to non-reimbursement for those charges generated solely because of participation in a research protocol.

NIMH has not articulated a formal position on this question, but clinical care in the three large trials mentioned earlier, including most medications, is routinely reimbursed under standard insurance mechanisms. However, blinded, placebo-controlled drugs and specific research assessments (in contrast to clinical assessments) typically are paid for by research funds. It would therefore appear that providers in practical clinical trials are legally
able to bill under standard insurance mechanisms, and, where blinding and use of placebo are not an issue, patients will be able to use standard prescription benefits to purchase medication. In addition, because no additional charges will be added to the cost of usual care as a function of participating in research, it is anticipated that there will be no economic or legal constraints to adding research procedures to standard clinical care.

Clinical psychopharmacology has been and likely will remain heavily influenced, if not dominated by, the pharmaceutical industry, especially for compounds early in the product development sequence. Industry funding for clinical trials is many times larger than NIH (extramural, including NIMH) funding: $4.1 billion, compared to $850 million in 2000, with only 10% of industry funds devoted to phase-IV trials (65). Because most if not all practical clinical trials would be classified as phase-IIIb or phase-IV trials, these trials would be disadvantaged with respect to potential funding either by industry or NIH. It is important to note that although industry-sponsored research is critical to new product development, its emphasis is on meeting U.S. Food and Drug Administration (FDA) regulatory requirements and on obtaining expanded marketing claims, not on evaluating the effectiveness of products as used in the general population. As a result, industry-sponsored research often fails to address broad public health needs or the needs of individual practitioners seeking to make good clinical decisions for individual patients. This shortcoming of industry-sponsored research is especially pertinent for decisions regarding risk, use of adjunctive treatments to improve partial response, maintenance and discontinuation of treatment, and transportability of treatments from the research to the clinical setting.

Although a compelling scientific argument can be made for practical clinical trials funded by industry (fewer negative findings and more definitive answers to safety questions, among other reasons), it is unlikely, although not inevitably so, that pharmaceutical companies will pursue a practical clinical trials agenda in psychiatry if doing so will, in their perception, put profits at risk, even when the answers would be of substantial public health importance. Particularly when comparing newer to older off-patent treatments, the risk of an adverse outcome (including a true tie) would be too great. Hence, in contrast to other areas of medicine where practical clinical trials are the standard for phase-III trials (10), industry-funded practical clinical trials in psychiatry in the short-term future will likely emphasize treatase addition studies—as in the epilepsy studies model in which a newer drug is added to an older compound to incrementally improve overall outcome (66)—if only because the requisite subject numbers are more readily available in the practical clinical trial framework. However, it may be worthwhile to point out that a more mature view of the role of regulatory agencies in “anticipating the market” would be to manage date that practical clinical trials should be done from the very beginning for psychotropic medications that will have widespread dissemination to the public. If the clinicians who prescribe these drugs would demand practical outcomes trials and, not incidentally, insist on the ethical obligation to make the results of human experiments public, the pharmaceutical companies would be more likely to do practical clinical trials and to publish the results.

Summary and Conclusions

After decades of industry-sponsored efficacy studies, psychiatry appears primed to follow the rest of medicine in moving toward practical clinical trials that seek to identify clinically relevant treatment outcomes of substantial public health importance. Such trials are necessary to evaluate clinically important interventions in a diverse population of study participants recruited from heterogeneous practice settings by using outcomes data that cover a broad range of ecologically valid health indicators (9). Without practical clinical trials, the heterogeneity in practice that yields regrettable heterogeneity in quality of care will continue unopposed because of the all-too-limited and often irrelevant evidence base (50). To implement the practical clinical trials model in psychiatry will require network construction and maintenance, procedures for matching trial aims to clinical practice requirements, best-clinical-practice assessments and treatments, selection of appropriate control or comparison conditions, simple yet reliable diagnoses and endpoints, feasible quality assurance procedures, and innovative data collection, management, and analytic procedures. For these developments to happen, a range of stakeholders, including NIMH, FDA, the clinical research community, and consumer advocacy groups, must work together to make practical clinical trials in psychiatry a priority.

References


Received Nov. 21, 2003; revisions received April 19 and June 12, 2004; accepted July 8, 2004. From the Department of Psychiatry and Behavioral Sciences and Duke Clinical Research Institute, Duke University Medical Center, Durham, N.C. Address correspondence and reprint requests to Dr. March, Department of Psychiatry and Behavioral Sciences, Duke Child and Family Study Center, 718 Rutherford St., Durham, NC 27705; jsmarch@acpub.duke.edu (e-mail).

Supported by NIMH grants 1 K24 MH-01557 and 1 P30 MH-66386 and by NIMH contract 9B-DS-0008 to Dr. March.
27. Weisz JR: Agenda for child and adolescent psychotherapy research: on the need to put science into practice (comment). Arch Gen Psychiatry 2000; 57:837–838
29. Weisz JR: Agenda for child and adolescent psychotherapy research: on the need to put science into practice (comment). Arch Gen Psychiatry 2000; 57:837–838
Korn A, Rimoin D: Central challenges facing the national clinical research enterprise. JAMA 2003; 289:1278–1287
64. Bridging Science and Service: A Report by the National Advisory Mental Health Council’s Clinical Treatment and Services Research Workgroup. Rockville, Md, National Institute of Mental Health, 1999
66. Perucca E, Tomson T: Monotherapy trials with the new antiepileptic drugs: study designs, practical relevance and ethical implications. Epilepsy Res 1999; 33:247–262
Somatoform Disorders: 
Time for a New Approach in DSM-V

Laurence J. Kirmayer, M.D., F.R.C.P.(C.)
Greg Simon, M.D., M.P.H.
Kurt Kroenke, M.D.

Objective: DSM-III introduced somatoform disorders as a speculative diagnostic category for somatic symptoms “not explained by a general medical condition.” Although retained and enlarged in DSM-IV, somatoform disorders have been the subject of continuing criticism by both professionals and patients. The extended period of preparation for DSM-V offers an important opportunity to reconsider the category of somatoform disorders.

Method: Exploration of the diverse aims of a diagnostic classification indicates that the authors must not only address the conceptual and practical problems associated with this category but also reconcile it with the parallel medical descriptive classification of functional symptoms and syndromes.

Results: The existing somatoform disorders categories require modification. The authors favor the radical option of the abolition of the categories. Diagnoses currently within somatoform disorders could be redistributed into other groupings, and the disorders currently defined solely by somatic symptoms could be placed on axis III as “functional somatic symptoms and syndromes.” Greater use could be made of “psychological factors affecting medical condition” on axis I. The authors suggest supplementing the diagnosis of functional somatic symptoms with a multiaxial formulation.

Conclusions: The authors promote a classification of somatic symptoms in DSM-V that is compatible with that used in general medicine and offers new opportunities both for research into the etiology and treatment of symptoms and for the greater integration of psychiatry into general medical practice.

DSM-III introduced the somatoform disorders as a speculative diagnostic category. It remained in the successive versions of DSM-III-R and DSM-IV. Although it has succeeded in focusing attention onto previously neglected patients with somatic symptoms that are unexplained by a general medical condition, we argue that it has failed in its declared purposes of aiding understanding, guiding research, and providing a useful basis for treating these patients (1). Since the planning process for DSM-V will allow a prolonged period for discussion and research (2), there is now an important opportunity to reconsider the somatoform disorder category. We propose a radical option: the abolition of the somatoform disorders as a category and the use of axis III to code somatic symptoms. Although this proposal is undoubtedly controversial, we hope that our arguments will both stimulate a more radical debate about how somatic symptoms are classified in DSM and inform the wider discussion about the general principles of the psychiatric diagnostic classification.

The Clinical Problem

Our starting point is the clinical problem (3). It is important to note that this is not just the small number of patients who come to psychiatrists with somatic complaints but the much larger number who are seen by all types of doctors with somatic symptoms that are not well explained by general medical conditions. Such symptoms account for a quarter to a half of presentations in both primary and secondary care (4). Although frequently minor or transient, these symptoms are often associated with a degree of distress and disability sufficient for them to be legitimately regarded as illnesses (5). Despite the size and importance of this problem, medicine—especially Western medicine—has found these conditions difficult to name, conceptualize, and classify (6). The names proposed have been bewildering in their variety and include somatization, somatoform disorders, medically unexplained symptoms, and functional symptoms. Conceptually, these illnesses lie in an ambiguous area of medical thinking somewhere between medicine and psychiatry (6). Their classification reflects this confusion: in psychiatry, they are classified as somatoform disorders (DSM-IV) and in medicine as functional somatic syndromes (5).

Current Psychiatric Terminology and Classification

The abolition of neurosis in DSM-III and its replacement by multiple new diagnoses led to acerbic debate (7). In this context, somatic syndromes that were neither ex-
plained by a general medical condition nor clearly associated with depressive or anxiety diagnoses were combined to create a new category of somatoform disorders. Central to the category of somatoform disorders was the newly proposed diagnosis of somatization disorder. Also included was a disparate group with other diagnoses, united only by their presentation with somatic symptoms. These diagnoses were conversion disorder, hypochondriasis, and psychogenic pain disorder. In addition, there was a residual category of atypical somatoform disorder.

In the subsequent revisions of DSM-III as DSM-III-R and DSM-IV, minor changes to the definitions of these disorders were made. There was also one major change—the introduction of undifferentiated somatoform disorder. The addition of this new diagnosis was necessary to provide a home for the large number of patients who, although clearly ill, did not fall within the existing somatoform categories. As a result, the category of somatoform disorders changed between DSM-III and DSM-IV from being a small grouping of relatively uncommon conditions to a general category covering a wide range of illnesses. The somatoform disorders currently listed in DSM-IV are shown in Table 1. ICD-10 was developed in parallel with DSM-IV and also includes a similar, but not identical, category of somatoform disorders embedded within the broader “neurotic, stress-related, and somatoform disorders” category (Table 1).

This review can be seen as part of a wider debate about the principles of psychiatric taxonomy that should be adopted for DSM-V. The issues under discussion include the relative merits of categorical versus dimensional approaches and the value of descriptive versus etiological classifications, the importance of utility as well as validity and reliability, the thresholds chosen for caseness, and the role of the social impairment criteria in defining caseness (9, 10). While mindful of this wider debate, this article will focus on issues specific to somatoform disorders.

The Purposes of Diagnosis

We first need to consider the main purpose of psychiatric diagnosis. In theory, this is to provide names for and categories of illnesses to aid communication, provide prognostic information, and guide treatment and research (11). In practice, however, diagnostic classifications have other functions that vary according to the user:

- Psychiatrists and other “mental” health specialists, especially those working in general medical settings, are often called upon to diagnose and treat patients with more severe and persistent somatic symptoms. They need diagnoses that perform the functions mentioned that also justify these conditions as appropriate for psychiatric attention.
- Patients are not the passive recipients of diagnoses; they also have expectations of a diagnostic label. It must be acceptable to them, appropriately represent their experience of suffering, imply a plausible explanation of what is wrong with them, and preferably lead to effective treatment. Diagnoses also have important implications for their social responsibilities and help them determine their expectations of health care and disability payments.
- Primary care practitioners see and manage the large majority of these patients (4). Although their primary task is to identify the symptoms that indicate serious and life-threatening medical conditions, they also have to describe and manage patients whose somatic symptoms are not associated with pathology. They need a simple and usable classification for this purpose.
- Employers, lawyers, insurers, those responsible for health benefits, and health planners all need a workable language and diagnostic system for all medical presentations associated with disability, health, and social costs, including those symptomatic presentations.

Shortcomings of Somatoform Disorders as Diagnoses

Although it is unrealistic to expect a diagnostic classification to meet all the demands that may be placed on it, many clinicians believe that the current terminology and classification system performs poorly in respect to almost all of the functions of diagnosis just listed.

Shortcomings of the Somatoform Category

1. The terminology is unacceptable to patients. With increasing transparency in health care, the acceptability of diagnostic terms to patients is important. Although proposed as an atheoretical term, “somatoform” is

| Table 1. DSM-IV Codes and Categories for Somatoform Disorders and ICD-10 Equivalents |
|---------------------------------|---------------------------------|
| **DSM-IV** | **ICD-10** |
| Code | Category | Code | Category |
| 300.81 | Somatization disorder | F45.0 | Somatization disorder |
| 300.81 | Undifferentiated somatoform disorder | F45.1 | Undifferentiated somatoform disorder |
| 300.11 | Conversion disorder | F44.0 | Dissociative (conversion) disorders |
| 307.80 | Pain disorder | F45.4 | Persistent somatoform pain disorder |
| 300.7 | Hypochondriasis | F45.2 | Hypochondriacal disorder |
| 300.7 | Body dysmorphic disorder | F45.3 | Somatoform autonomic dysfunction |
| 300.81 | Somatoform disorder not otherwise specified | F45.8 | Other somatoform disorders |
| 300.81 | Somatoform disorder not otherwise specified | F45.9 | Somatoform disorder unspecified |
| 300.81 | Somatoform disorder not otherwise specified | F48.0 | Neurotic disorders (in other neurotic disorders category) |
commonly seen as related to the older term “somatization” (12). This implies that the symptoms are a “mental disorder” in somatic form and may be regarded by patients as conveying doubt about the reality and genuineness of their suffering (13).

2. The category is inherently dualistic. The idea that somatic symptoms can be divided into those that reflect disease and those that are psychogenic is theoretically questionable (6). Indeed, the view that symptoms can be “explained” solely by a disease is a debatable one that is not entirely in accordance with empirical data (14). In practice, most physicians adopt a broad perspective when assessing a patient’s symptoms (15).

3. Somatoform disorders do not form a coherent category. The only common feature of somatoform disorders is that they show somatic symptoms without an associated general medical condition. Beyond that, they lack coherence (9). The overlap with the many other psychiatric disorders that are also defined in part by somatic symptoms, such as depression and anxiety, is also a potential cause of misdiagnosis.

4. Somatoform disorders are incompatible with other cultures. Somatoform disorder diagnoses do not translate well into cultures that have a less dualistic view of mind and body (for example, the current Chinese classification is based on DSM but specifically excludes the somatoform disorder category (16)). Exporting a dualistic diagnosis of somatoform disorder to these cultures at the same time that Western medicine is trying to escape it would seem to be counterproductive.

5. There is ambiguity in the stated exclusion criteria. The diagnosis of somatoform disorder requires the exclusion of general medical conditions. However, there is lack of clarity about which medical diagnoses should be regarded as exclusionary: for example, do medical “functional syndromes,” such as irritable bowel syndrome, count as exclusions? One consequence of this lack of clarity is that patients may be classified as having both an axis III disorder (for example, irritable bowel syndrome) and an axis I somatoform disorder (such as undifferentiated somatoform disorder or pain disorder) for the very same somatic symptoms. This seems to be ridiculous.

6. The subcategories are unreliable. Many of the subcategories of somatoform disorders have failed to achieve established standards of reliability (17).

7. Somatoform disorders lack clearly defined thresholds. The lack of any clearly defined threshold for what merits a somatoform disorder diagnosis has led to disagreement about the scope of this category and also to its gradual enlargement (18). It is probably for this reason that most major epidemiological surveys of psychiatric disorders have excluded somatoform disorders.

8. Somatoform disorders cause confusion in disputes over medical-legal and insurance entitlements. Somatoform disorder diagnoses have proved problematic in relation to medical-legal and social security entitlements. On one hand, they can provide spurious diagnostic validation for simple symptom complaints, and on the other, they can undermine the reality of somatic symptoms as “merely psychiatric.” They thereby provide considerable scope for generating irresolvable differences of opinion.

In summary, the existing category of somatoform disorders may be regarded to have failed.

Shortcomings of the Specific Somatoform Subcategories

Somatization disorder is arguably the archetypical diagnosis of the somatoform disorder category. Its introduction was influenced by the then-recent work of the St. Louis group (8). Arguably, it has subsequently received attention out of proportion to its prevalence relative to that of the other somatoform disorders. Furthermore, doubts have been expressed about both its clinical value and conceptual basis (19). First, patients with somatization disorder have prominent psychological as well as somatic symptoms so that the syndrome is hardly an exemplar of a predominately somatic condition (20). Second, it has a substantial overlap with personality disorders, particularly borderline personality disorder (21). Third, although the requirements for diagnosis are unusual in that they rely on a lifetime history of symptoms, there is evidence that patients’ recall of past symptoms is variable and that the diagnosis has low stability in longitudinal surveys (22). Fourth, it is based merely on counting the number of “unexplained” somatic symptoms and so lacks even face validity as a psychiatric disorder. The number of somatic symptoms a person reports is continuously distributed in the general population, and the diagnosis merely represents an extreme of severity on what appears to be a continuum of distress (23). Finally, the diagnosis of somatization disorder offers the practitioner little specific guidance about treatment beyond clinical management aimed at minimizing health care use and iatrogenic illness (24).

In response to the observation that many patients with chronic multiple symptoms do not meet the DSM-IV criteria for somatization disorder, attempts have been made to reduce the number of symptoms required for a diagnosis (25, 26). Although these proposals have the advantage of acknowledging that the number of somatic symptoms forms a continuum, they retain the limitations of a diagnosis based almost exclusively on simply counting somatic symptoms.

Hypochondriasis as a diagnostic category remains controversial. Although there is good evidence of the co-occurrence of the triad of disease conviction, associated distress, and medical help-seeking, these symptoms are arguably better conceived of as a form of anxiety that hap-
Conversion disorder has long been a problem for diagnostic classification. DSM-III placed it with other diagnoses in the somatoform section because of the shared characteristic of somatic symptoms that are not intentionally produced (7). The DSM-IV workgroup recognized a close relationship with dissociative disorders but confirmed the DSM-III classification (29). We propose that this discussion should be revisited.

Body dysmorphic disorder remains uncomfortably placed in the somatoform disorder category. There have been persuasive arguments that it should be rehoused; in particular, the suggestions that it might be better grouped with obsessive-compulsive disorder (30) could be usefully revisited.

Undifferentiated somatoform disorder was placed alongside somatoform disorder not otherwise specified (the successor to atypical somatoform disorder) in DSM-III-R as a poorly defined catchall for the patients who did not fit into the original specific DSM-III categories (31). However, it soon became clear that these were not merely small residual diagnoses but rather the most widely applicable categories. Even though this diagnosis is not widely used in clinical practice, its existence represents the need to have a diagnosis for a very large group of patients not easily classified elsewhere.

Pain disorder has undergone significant revision between DSM-III and DSM-IV. However, as noted by the Working Group for DSM-IV, there remain problems both in its definition and in establishing it as a separate disorder (32).

In summary, many of the diagnostic subcategories currently housed within the somatoform disorders either lack validity as separate conditions or could be better housed elsewhere.

The Somatoform Classification and the Evidence

We need to consider the compatibility of DSM-IV somatoform disorder with existing evidence.

Are Somatoform Disorders Consistent With Evidence About Epidemiology?

Studies of primary care patients (22, 31, 33–35) have repeatedly found that the core DSM- or ICD-defined somatoform syndromes (such as somatization disorder or hypochondriasis) are relatively rare, whereas the more vaguely defined but often clinically important diagnoses are common. This finding makes the existing classification of limited value. Measures of disease impact (such as disability or comorbid psychiatric disorder) are consistent with the apparently well-defined somatoform disorders being better regarded as the extreme on a continuum of illness (23). Population-based research also provides only limited support for the particular syndromal patterns of symptoms described by the somatoform categories (36). Cross-sectional studies report that all types of somatic symptoms (whether explained or unexplained by identifiable disease) are associated with symptoms of anxiety and depression (31, 35, 37). Furthermore, longitudinal studies have found that the type of symptoms patients report frequently varies over time (33). All of these findings raise doubts about both the validity and utility of existing somatoform disorder diagnoses.

Are Somatoform Disorders Consistent With Evidence About Etiology?

The somatoform criteria and their accompanying text are based largely on the etiological concept of “somatization,” a hypothetical process whereby mental illness manifests as somatic symptoms. Modern evidence suggests that this conceptualization is simplistic; it favors instead a multifactorial etiology with interacting psychological, social, and biological factors (38) (Figure 1). It is especially important to note that there is increasing evidence that biological factors are relevant (6). Other factors that influence symptoms include their modulation by depression and anxiety (39); processes of perception and symptoms interpretation (40); the reactions of other people (family,
friends, acquaintances) (41); and iatrogenic processes, as well as the influence of the insurance, compensation, and disability systems (42).

**Is the Existing Somatoform Diagnostic Classification Workable in Clinical Practice?**

We have found the existing somatoform disorder category problematic in practice. The first problem is diagnostic confusion resulting from the overlap of somatoform disorders with the classification used by practitioners of internal medicine. The latter defines “functional syndromes” descriptively, according to either the patient’s major symptom (for example, dizziness, tension headache) or the bodily system that these symptoms appear to be associated with (for example, noncardiac chest pain, irritable bowel syndrome) and provides an alternative diagnosis to somatoform disorder for the same patient (5). The second problem is whether somatoform disorder can be diagnosed in patients who also have a general medical diagnosis. For example, atypical noncardiac chest pains are common and often distressing in those who have suffered myocardial infarction or undergone cardiac surgery. Should they also be given a diagnosis of somatoform disorder? The third problem relates to the conventional restriction of the diagnosis “psychological factor affecting medical condition” to patients who also have a general medical diagnosis. If a patient is distressed about somatic symptoms, should the diagnosis be somatoform disorder or psychological factor affecting a medical condition and should that depend on whether the symptoms are considered to be general medical or psychiatric in nature? In our view, the question of whether a condition is regarded as medical or as psychiatric is not an indication of etiology but simply a pragmatic statement about which medical specialty is the best place to manage it, in the same way that some conditions may be considered medical and others surgical. With these criteria, somatoform disorders may be considered to be as much general medical as psychiatric conditions.

**Approaches to Change**

How might we improve the current somatoform disorder classification? As a first step, we examine the features of the existing medical and psychiatric diagnostic systems that have proved to be useful and that might be incorporated into a new approach.

**What Can We Learn From the Nonpsychiatric Medical Approach?**

Because most patients with somatic symptoms unexplained by general medical conditions are seen by nonpsychiatric medical practitioners, we might usefully consider how these doctors currently classify them. The “nonpsychiatric” terminology and classification as functional syndromes have advantages. In particular, simple symptom descriptors, often qualified with the term “functional,” have the advantage of being atheoretical. The term “functional,” although sometimes used as a code for “psychological,” originally meant a disturbance of function as opposed to structure (43). In this sense, it is usefully nondualistic and is compatible with recent research findings that indicate altered physiological function in many of these conditions (6). It is also acceptable to patients (13). Similarly, specific functional syndrome labels, such as irritable bowel syndrome and fibromyalgia, also have face validity and appear to be largely acceptable to patients. Although there is evidence for overlap (or comorbidity) among these specific functional syndromes (5), current treatment approaches generally target specific symptom clusters (such as those of irritable bowel syndrome), and simple symptoms descriptors, perhaps qualified with by the term “functional” to indicate the absence of pathologically defined disease, could usefully therefore be retained in a new classification.

**What Can We Learn From the Existing Psychiatric Approach?**

While we have highlighted the shortcomings of the current DSM classification in describing somatic symptoms, DSM-IV does have useful properties that could be further exploited. It has highlighted that groups of illnesses of somatic symptoms are unexplained by general medical conditions and has offered a system for classifying them. It has transcended the bodily-system-based approach of defining functional syndromes that has been shown to be of limited validity (5). Furthermore, the multiaxial nature of the system helpfully allows the separate coding of medical and psychiatric diagnoses and provides more useful information about patients than a single diagnosis.

**What Should Be the Properties of the New Classification System?**

We propose that any new system of classification should

- Be consistent with the general classification principles of DSM-V
- Relate effectively to the functional disorder classification used by physicians
- Be acceptable to patients
- Be etiologically neutral
- Be helpful in planning treatment
- Be equally applicable to patients with general medical conditions
- Provide an effective basis for further research

**A Proposal for DSM-V**

A conservative option would be a simple revision of the categories and definitions of somatoform disorder, accompanied by a rewriting of the accompanying text to take account of the issues outlined. This approach would, however, be similar to that taken by the DSM-IV work-
The suggested revision, renaming, and redistribution of the existing somatoform categories are described here and shown in Table 2. We also argue that more use could also be made of the category “psychological factors affecting medical condition” that appears in DSM-IV only as part of the chapter titled “other conditions that may be a focus of clinical attention.” This could be an axis I accompaniment to any axis III diagnosis.

- Somatic symptoms associated with depression are classified with depression and those with anxiety with anxiety (with an additional specification “with prominent somatic symptoms” to reflect the patient’s preoccupation with—or concern about—somatic symptoms, such as fatigue, pain, or physical malaise).
- Hypochondriasis should be renamed as health anxiety disorder and placed within the anxiety disorders. Although it overlaps with other forms of anxiety, the focus on disease or medical diagnosis is clinically important and influences presentation and treatment (44). “Illness fear” probably fits best within phobias (45).
- Body dysmorphic disorder has never sat comfortably in the somatoform disorders category and should be moved elsewhere, possibly together with obsessive-compulsive disorder (30).
- The classification of dissociative and conversion symptoms requires review. One option would be to place them together as a separate subgroup defined by criteria similar to those in current use (29).
- The most difficult problem is the classification of the continuum of conditions defined merely by a number of somatic symptoms; this ranges from somatization disorder to undifferentiated somatoform disorder. We suggest that these symptoms are classified on axis III as “somatic symptoms” or as “functional somatic symptoms.” Associated psychiatric diagnoses could be coded on axis I with other factors justifying psychiatric intervention as “psychological factors affecting a medical condition.” Hence, the approach would be the same for a patient with irritable bowel syndrome as for a patient with illness worry after coronary artery surgery. It remains possible that it will eventually be possible to propose convincing additional psychiatric diagnoses defined by abnormalities in psychological or behavioral processes to replace this category, but this requires further research.

Axis II

- Patients with personality disorders as well as somatic symptoms will have their disorder coded on axis II as before.
- Somatization disorder may be better regarded as a combination of personality disorder (axis II) with affective or anxiety disorder (axis I) (46).

### TABLE 2. A New Classification for Conditions Referred to in DSM-IV as Somatoform Disorders

<table>
<thead>
<tr>
<th>Axis I (psychiatric diagnoses)</th>
<th>Axis II (personality disorder)</th>
<th>Axis III (medical conditions)</th>
<th>Axis IV (social and environmental problems, including interaction with the health care system)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenia and delusional disorders</td>
<td>Personality disorder</td>
<td>Functional somatic symptoms and syndromes</td>
<td>Pain and fatigue, physical malaise and concern about or as functional somatic symptoms</td>
</tr>
<tr>
<td>Depressive disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety disorders, including anxiety/hypochondriasis and specific phobia (illness fears)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversion/dissociation disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological factors affecting general medical condition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axis III diagnosis (medical conditions), including (functional) somatic symptoms and syndromes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axis IV psychosocial and environmental problems, including interaction with the health care system</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

groups and could consequently be expected to lead to similar difficulties. A more radical option not open to the revisers of DSM-IV would be the abolition of the category of somatoform disorders altogether, with reassignment of the specific somatoform diagnoses to other parts of the classification. We favor the second option. Our specific proposals are to

1. **Abolish the somatoform disorder category.**

   The somatoform disorder term, concept, and category have failed psychiatrists, nonpsychiatric physicians, and patients. There seems to be little reason to retain them.

2. **Adopt a new term for somatic symptoms and syndromes.**

   An alternative term to “somatoform” is needed to describe symptoms, especially those that are not closely related to a general medical condition. These could just be called “somatic symptoms” with an associated disease diagnosis specified when appropriate (e.g., “pain” and “pain associated with lung cancer”). If a general adjective is required to emphasize the lack of association with a general medical condition, we suggest “functional” in its original use as a strong candidate.

3. **Redistribute disorders currently listed under somatoform disorders in other parts of the classification.**

   This task has several aspects. First, several disorders currently housed in somatoform disorders can simply be moved to other axis I (psychiatric disorders) or axis II (personality disorder) categories. Second, it could be clarified that the axis III (general medical conditions) label is to be used for all those somatic symptoms most commonly managed by general medical doctors, regardless of whether the patient has a disease diagnosis. Third, axis IV may be used to describe unhelpful interactions with medical services, as well as access to them. Specific proposals for each axis are as follows:

   **Axis I.** The axis I classification needs modification so that the specific diagnoses currently within somatoform disorders are either redistributed elsewhere or reformulated.

   **Axis II.**

   - Patients with personality disorders as well as somatic symptoms will have their disorder coded on axis II as before.
   - Somatization disorder may be better regarded as a combination of personality disorder (axis II) with affective or anxiety disorder (axis I) (46).
Axis III

- Somatic symptoms and syndromes and pain disorder could be classified on axis III. This should be seen as no more than as a clarification of the existing DSM-IV principle: “Axis III is for reporting current general medical conditions that are potentially relevant to the understanding or management of an individual’s mental disorder.” These conditions are classified outside the “Mental Disorders” chapter of ICD-9-CM (and outside chapter V of ICD-10).

- There would be advantages in establishing mild, moderate, and severe levels of somatic symptoms with current DSM-IV severity criteria with a threshold set so as to be of clinical significance. Somatic symptoms are defined in DSM-IV as “causes of clinically significant distress or impairment in social, occupational, or other important areas of functioning.”

Axis IV. Axis IV currently allows a listing of psychosocial and environmental problems. It could be usefully extended to include unhelpful interactions with the medical care system, such as frequent attendance (47).

4. Elaborate on diagnoses with an additional multidimensional description.

We also argue that the new diagnostic classification would benefit from a supplementary description of individual patients in terms of descriptive dimensions as well as a diagnosis (38). An additional multidimensional classification would be consistent with our current etiological understanding. This would be valuable for more clearly defining patients to be included in research and for targeting treatment interventions at specific etiological factors. It might be reproduced as an annex to DSM-V. There are already analogies in the DSM-IV sections on pain and sleep in reference to other, more detailed classifications. A suggested scheme is shown in Table 3. This multifactorial approach could profitably be further developed and publicized not only within psychiatry but also within medicine as a whole.

Implications of the Proposals

It can be expected that the DSM-V manual will, like its predecessors, become a standard text. It will therefore have an important role in improving the general understanding of those conditions currently classified as somatoform disorders. We anticipate that if our suggestions are incorporated into DSM-V, they will have many positive implications for practice, research, and the wider understanding of these conditions. Although it would be unrealistic to expect that our proposed revisions will entirely eliminate unhelpful dualist thinking, we believe that they represent a useful step in that direction. We acknowledge that further debate and evaluation are required to adequately evaluate these proposals and to compare them with the alternative of a more limited revision of the current categories.

Clinical Implications

The main implication of our proposals is the acceptance of etiological neutrality about those somatic symptoms that are not clearly associated with a general medical condition. We propose a pragmatic classification that is explicitly based on the branch of medicine most concerned with the management of the condition rather than on presumed etiology; this has the consequence of emphasizing that some patients require attention from both general medicine and psychiatry. We anticipate that this would help integrate psychiatry with other medical specialties. It also offers a terminology and context that are more likely to be acceptable to patients and therefore more effective in engaging them in treatment. We already have treatments for functional somatic symptoms (48, 49), but we need better and less overtly “psychiatric” ways of explaining and implementing them.

Research Implications

The approach we have outlined is consistent with the evidence base and also indicates directions for further study. This research should ultimately lead to a more empirically based nosology with clearer implications for treatment. It should also clarify the etiological processes, including neurobiological, perceptual, cognitive, and behavioral factors, that underpin all symptoms. Epidemiological research could start with further analysis of existing data, for example, an examination of natural clustering and stability of syndromes over time. Finally, the supplementary multidimensional description will allow a more precise study of etiological factors and their evaluation in homogeneous groups of patients.

Wider Implications

Improved collaboration by those involved in the creation of psychiatric classifications with their nonpsychiatric counterparts will have wider benefits in encouraging a more integrated perspective on symptoms. This nonspecialty-based approach has been called symptoms research (50). Ultimately, we anticipate a merging of the classifications used by general medical and psychiatric physicians. This, in turn, will enhance the role of psychiatry in general medical care to the benefit of patients.

---

TABLE 3. A Multidimensional Descriptive System for Somatic Symptoms

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Type of somatic symptom</th>
<th>Number of symptoms</th>
<th>Course (e.g., acute, chronic, recurrent)</th>
<th>Disease pathology/pathophysiology</th>
<th>Health beliefs</th>
<th>Illness behavior</th>
<th>Associated psychiatric disorder</th>
<th>Social factors (e.g., employment, social benefits)</th>
</tr>
</thead>
</table>

http://ajp.psychiatryonline.org
References

20. Wetzel RD, Guze SB, Cloninger CR, Martin RL: Briquet's syndrome (hysteria) is both a somatiform and a "psychoform" illness: a Minnesota Multiphasic Personality Inventory study. Psychosom Med 1994; 56:564–569
The hippocampus is one of several CNS structures important for learning and memory. Its critical role in memory was first noted in a patient who lost the capacity for forming new memories after a bilateral hippocampectomy for uncontrollable seizures. A delineation of the exact role of the hippocampus in learning and memory has been pursued ever since. Laboratory rodents are good subjects for experiments in learning and memory because they learn new knowledge quickly and reliably in a fashion similar to humans. Different aspects of memory have been associated not only with frontal, hippocampal, and striatal structures, but also with the specific hippocampal subfields and their cortical layers. Recently developed transgenic techniques allow very selective molecular lesions to be made within the hippocampus to answer specific questions about the regions and transmitters involved in memory formation. The figure shows the hippocampus of a transgenic mouse, stained for a single molecule (NR1 mRNA) that directs the expression of this essential subunit of the NMDA receptor protein. Note that on the left side, there is an absence of NR1 mRNA (in the CA1 field of the hippocampus) compared with the right side. The area lacking the molecule is enlarged and stained with a nuclear stain to show there is no cell damage. Such a deletion of the essential subunit of the NMDA-sensitive glutamate receptor blocks information transmission at that receptor in that region. The molecular lesion is accomplished using a genetically engineered mouse that has inserted nuclear tags placed around the gene of interest (called loxP nucleotide sequences). A viral vector (an engineered adeno-associated virus that expresses Cre recombinase) is used to induce the removal of the tagged gene but only in the region where the viral vector is injected. This molecular approach is more selective and precise than other mechanisms of blockade, whether by pharmacological approach or by nonconditional knockouts. When we explored the behavioral consequences of such a knockout in CA3, we found that this lesion totally inhibits the ability of an animal to quickly learn a new set of facts within a specific context, but it does not interfere with the animal’s use of information already learned. These animal data suggest that CA3 in the human hippocampus contributes to our human ability to quickly gain a familiarity with a new set of facts relating to a given context and to remember them as related in this manner. This may be particularly important in social contexts to learn that some interpersonal responses are appropriate in particular contexts but inappropriate in others.

ROBERT W. GREENE, M.D., PH.D.
Dallas, Tex.

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).
How Many Psychiatrists Does It Take?

The dinner was excellent: gently seared tuna and succulent crab cakes. The symposium was not as enjoyable, but still, this was a good start to my stay in Philadelphia. I had flown in that morning for the annual meeting of the American Psychiatric Association, my first experience at such a large conference. My room for the week was at a decidedly seedy motel in Bucks County, about an hour away from the city. I looked at my watch. There was still some time before the last train back, so I decided to spend some more time at this hotel, one of Philadelphia’s finest. I walked toward the bar, situated in the center of a high-ceilinged lobby with shimmering chandeliers. An ingenious waterfall with thin sheets of water cascaded onto pebbles, making a pleasant sound.

I settled into one corner of the bar and ordered a cocktail, soaking in the ambience. Most days, this was probably an oasis of jazz music and soft murmurs, but tonight it was teeming with other psychiatrists. Like me, they had bags with the conference logo displayed prominently and badges. Unlike me, most of them were in groups engaged in animated conversation. I entertained myself by listening in on some of the discussion.

“The schizophrenia update was excellent. It’s going to definitely change my practice.”
“I am not so sure about the applicability of that data in the real world.”
“What’s your experience with…”

Just then, a man walked into the lobby, carefully taking the few steps to the bar. He sat down on a sofa next to me, and I could smell him from where I was sitting: sweat, urine, and stale cigarettes. His shirt was wet even though it hadn’t been raining, and he rubbed his unshaven face and looked around at the crowd, torn jeans sagging below a corpulent belly. He began mumbling to himself, and although I briefly entertained the idea that he might be an eccentric colleague, a fellow psychiatrist who had forsaken formal attire in favor of something more unconventional, he looked so destitute and beaten that it was obvious he was one of the many mentally ill and homeless people I had seen in the city. Drugs, alcoholism, a major mental illness, social drift, and here he was, rummaging through the torn plastic bag he was carrying.

It struck me as incredibly ironic that this should happen, that someone who was mentally ill should walk into a roomful of psychiatrists. By now, the bar personnel were staring at the man and whispering among themselves. One of them finally went to the man.

“Do you need anything?”
The man shook his head and mumbled, "No."
“This bar is for guests only, sir,” the waiter said.
The man ignored him and just looked away, saying something under his breath.

I found myself troubled by his presence, uncomfortable because I felt obliged to do something, and yet I held back, unsure of what I could do for the man. I wondered if my colleagues were similarly conflicted. They certainly seemed to have noticed him. I saw groups of other conference attendees pause and cast furtive but trained eyes on the man. The buzz of conversation around the bar slowed, and the soft strains of piano and the gentle splash of a waterfall could be heard again.

Then security guards appeared from either side of the lobby. With their suits and earpieces, they looked like they’d just taken time off from a presidential entourage. The man sank back into the sofa, but when they said, “Sir, you have to leave,” he collected his belongings and rose to his feet without a struggle.

“It struck me as incredibly ironic that this should happen, that someone who was mentally ill should walk into a roomful of psychiatrists.”
I sat there, sipping my cold martini, and convinced myself that there was little that anyone could do. The options were limited; psychiatric practice itself was inadequate, making interventions in situations like these impossible. I mean, we could hardly administer medications or perform psychotherapy right there in the room. What were we supposed to do? Now, if a man with chest pain walked into a congregation of cardiologists, that would be another story. There would be a scramble of cardiologists rushing to help the man—to administer advice, to recommend hospitalization, or, if the need arose, to start cardiopulmonary resuscitation. How bizarre if instead of helping the man with chest pain, the cardiologists ignored him and continued to discuss the latest data in their field. Or worse still, what if they felt uncomfortable—even embarrassed—that a man with chest pain had stumbled into a cardiology meeting? What if they just did not know what to do?

The man stood and looked around, eyes drifting from one side of the room to the other, and it seemed that he was seeking help. But I continued to sip my drink, and like every other psychiatrist in the room, I avoided eye contact with the man. He shambled off, making his way through the lobby, and pushed the double glass doors and went out into the cool Philadelphia night.

The room seemed to lighten when he left. Glasses clinked, laughter and merriment drowned out the music once again, and conversations about the latest treatments of mental illnesses resumed in full earnest.

SHYAM K. BHAT, M.D.

Address correspondence and reprint requests to Dr. Bhat, Department of Psychiatry and Internal Medicine, Southern Illinois University School of Medicine, P.O. Box 19636, Springfield, IL 62794-9642; sbhat2@siumed.edu (e-mail).
Dilemmas in the Psychotherapy of Sexually Impulsive Patients

Glen O. Gabbard, M.D.

Tanya J. Bennett, M.D.

Psychoanalytic therapy emerged from a sociohistorical era of repressed sexuality, and Sigmund Freud originally thought that “dammed-up libido” might be responsible for certain types of neurotic psychopathology. Much of the treatment was geared toward making patients consciously aware of their unconscious sexual wishes and the tenacity of their defenses against them. More than a century has passed since then, and the cultural context of sexuality in the early 21st century bears little resemblance to Victorian Vienna. Twelve-year-old girls emulate Britney Spears in their apparel. Teenagers are exposed to explicit sexuality on cable television and in the local cinema. Raw sexuality is ubiquitous in advertising, on magazine covers, and in television commercials.

Patients who come for psychodynamic/psychoanalytic therapy today are far less likely to be seen with sexual inhibitions or conflicts. Indeed, sexuality may be a relatively minor issue in their presenting clinical picture. The sexual preferences, fantasies, and practices of today’s patient may not even emerge until the process is well launched. Moreover, when sexual material does enter into the psychotherapeutic dialogue, it may be incidental to the main theme of the treatment and be entirely egosyntonic. The patient may express no interest whatsoever in changing sexual practices, and the therapist may need to accept the patient’s right to set the agenda for the therapy. These cases may present extraordinary challenges for the therapist because of their potential to cause harm to the patient.

Dr. Bennett

When I first met Ms. A, a 25-year-old woman, she was sobbing uncontrollably in the waiting room at the Baylor Psychiatry Clinic and was about 20 minutes late for our appointment. She had called the clinic earlier on her cell phone, reportedly lost and panicky, and had pleaded with clinic staff not to hang up until her arrival. In the interview room, Ms. A calmed down quickly and apologized for her tardiness and behavior. She appeared very assertive and remained dramatic even after she had regained her composure. I was immediately struck by her physical beauty, impeccable grooming, and stylish clothing. Her chief complaint was “my medication isn’t working,” and she stated that she had had depression and anxiety since age 10. She recounted numerous recent stressors, most notably her marriage a few months earlier. Her life had changed in many ways, including that she no longer needed to work or attend school, and she alluded to this transition as fairy-tale material, a dream come true. Despite this development, she endorsed multiple symptoms of depression and felt confused and purposeless. She ruminated about how horrible she was, sometimes for hours before falling asleep each night, and engaged in excessive apologizing. She complained of constant worry and expressed concern that her moods were up and down and out of control. She also described several episodes of what she called “paranoia” that involved people looking at her or following her—experiences that sounded credible rather than delusional.

Ms. A had been in psychotherapy for the presumed treatment of depressive symptoms with a well-respected psychiatrist and family friend from the ages of 13 to 16. During this time, she had not received any medication. She had had no subsequent psychiatric care until age 24, at which time she sought help for depression and anxiety, was given a prescription for an antidepressant, and saw a psychiatrist every 3 months or so for medication management.

Her pertinent medical history included very painful menstrual periods since menarche and monthly treatment-resistant vaginal infections since she became sexually active with her husband. She also reported a past history of polysubstance abuse, reportedly in full remission for 2 years.

Her developmental history was significant in that she was the third of three children born to first-generation immigrants. Ms. A’s mother was reportedly told that she could not get pregnant again after Ms. A’s birth. Ms. A described her mother as depressed, overly protective and manipulative, and chronically ill with multiple health problems. Her mother was also the daughter of two victims of severe trauma, and her maternal grandmother reportedly “went crazy” after her grandfather’s death. Ms. A’s father worked until late at night, and Ms. A described her childhood as spent primarily with her mother. She described herself as a rebellious, argumentsative adolescent who frequently fought with her mother, who would laugh openly at her the angrier Ms. A became and who would punish her by not speaking to her for days. During her school years, she felt that she did not belong and had few female friends but always wanted them. She attributed this to their being jealous but hesitated to link it to her beauty. She added, however, that despite everyone always telling her that she was beautiful, she didn’t know if she believed it or not.

“Her description of these events featured rooms full of naked people drinking, talking, and involved in various sexual acts, some behind closed doors and many not.”
When asked what her goals were in coming to the clinic, she stated that she wanted to get her medication adjusted and was interested in therapy. When asked what she would like to work on in therapy, she stated that her main concerns were her inability to show her husband how much she loved him and to “calm... down.” I diagnosed her with dysthymia, anxiety disorder not otherwise specified, and polysubstance abuse in full remission based on her history. I made a mental note of her histrionic traits, restarted her antidepressant, and scheduled our first therapy session for the following week.

Ms. A started our first therapy session by joking about a recent dream. In the dream, her husband and his sister were having sex in a room with a large plate-glass window. Her husband was fondling his sister’s nipples. Ms. A was outside watching them through the window. Her husband rose up partially from the missionary position in obvious enjoyment. He turned and looked at her and smiled. She did not attempt to work with the dream in therapy and seemed uninterested in its meaning. For the next 7 months or so of therapy, Ms. A would come in, sit down, and start talking, not stopping until the session was finished. As her dream suggested, she focused primarily on her feelings of isolation and being an outsider. She also felt that she did not know who she was anymore or what she wanted. She could not separate her own needs and desires from her husband’s. She grappled with the things rich people do, which she often viewed as shocking or wrong, and the demands of her new role as a wife in a well-known and socially visible family. She spoke “at” me rather than “to” me in a rapid and dramatic (but not pressured) manner and would often talk through my occasional comments or interrupt me before I had finished. If I did finish, she would ignore, disregard, or minimize my comment almost immediately. I felt angry, frustrated, and inadequate. I might as well not have even been in the room. The sessions felt completely beyond my control, and I anticipated them with dread. I was gradually becoming aware of a characterological style that made me wonder about borderline or histrionic traits, especially because her depressive and anxiety symptoms appeared context-dependent.

Dr. Gabbard

Ms. A’s first dream communicated important themes to Dr. Bennett while she simultaneously disavowed them as “just a dream.” In effect, she showed a compromise between telling and not telling her therapist about her sexual concerns. Because Ms. A did not provide extensive associations to the dream, we could not be certain of its meaning. However, the manifest content was highly suggestive. Ms. A was an observer watching her husband engaged in incestuous sexual relations with his sister. She was feeling excluded from something forbidden. She felt hurt and betrayed. The characters in a dream are often like the dramatic personae of a play, with each of them representing different aspects of the dreamer—both self and object representations that may be conscious or unconscious. Hence, one way to understand the dream is that Ms. A may also have been revealing to Dr. Bennett an important aspect of herself, a feeling of always being an excluded onlooker. Indeed, Dr. Bennett then told us that she, too, often felt like an outsider in the early months of therapy. As she put it, she might as well not have been in the room. Ms. A was beyond her control in the same way that, in the dream, Ms. A’s husband was out of her control. Hence, we can speculate that Ms. A may have recapitulated an internal drama depicted in the dream but with the roles assigned differently. By “performing” for Dr. Bennett, Ms. A was in the role of the sexual performer in the dream while Dr. Bennett was identifying with the part of the patient that felt like an outsider looking in. Some patients convey who they are not by a clear narrative of their inner world but by making the therapist feel as they do. In technical terms, we might refer to this phenomenon as a projective identification of a self-representation rather than as an objective representation from the patient’s inner world. This identification with the patient helps the therapist empathize with the patient’s experience. The patient’s jocular attitude may have expressed a need to distance herself from the dream by making fun of it with her therapist and dismissing its importance.

Dr. Bennett

After about 7 months into therapy, Ms. A came in one day and stated, “I haven’t been honest with you.... I haven’t told you about a big part of my life because I’m scared you won’t like me.” She then proceeded to tell me about her sex life, which included sadomasochistic sex with her husband as well as their “swinging” lifestyle. This was somewhat awkward because I did not know what “swinging” meant and said so. This obviously pleased Ms. A, who smiled and laughed a little, stating that now she could teach me something. She proceeded to define swinging as similar to dating but for couples rather than individuals. I simply listened as she described a meeting that she and her husband had with a couple for dinner and drinks. After dinner, they went to a hotel room together, where they talked some more, including about what they liked sexually, and then had sex. Ms. A stated that the sexual activity consisted of male-female or female-female kissing and touching, followed by intercourse with one’s spouse while the other couple watched. She stated that they would always talk about everything afterward and that she really liked this part of the experience. In terms of the sadomasochistic sex, this was something that she and her husband did alone together. She vaguely described him as the aggressor, downplaying it as all in fun—nothing serious—but did not supply any details at this time. Somewhat defensively, she stated that her husband believed that if it felt good, do it, and that what two or more consenting adults did privately was their own business and that they should not be judged negatively for it or considered freaks. She added that she agreed with him and that she loved sex, had always wanted lots of sex, and couldn’t get enough. She also talked about “wanting to know women” and about how she had never had any close female friends and craved this intimacy. Ms. A openly questioned her sexuality and the conventional societal definitions of heterosexuality versus homosexuality. She defined sexuality as more of a spectrum.
During this session, I felt like I was caught up in an embarrassingly riveting made-for-television movie. The content was titillating, and Ms. A was obviously aware of this and enjoyed the delivery. Her dramatic tone made it very easy to visualize the foursome. I also remember picturing the shock on my supervisor’s face as I told her about this session. I did not feel particularly concerned for Ms. A’s safety at the time. She had essentially described the swinging as talking, petting, and sexual intercourse only with her husband and had minimized the danger involved in the sadism and masochism. It felt like she had customized the script on my behalf—emphasizing the talking, openness, and honesty and downplaying the risks. Perhaps because of this, I did not feel repulsed by her actions or judge her negatively for them. It did not cross my mind that this was perverse or deviant but instead seemed more playful or exploratory. I found myself rationalizing the behaviors right along with her. In general, I agreed that, for the most part, what consenting adults did in private was indeed their business, as long as they were not hurting themselves or others. I also agreed that sexuality is not always black and white. Despite this, I felt uncomfortable and awkward. This intensified when she stated that several couples they had met while swinging were medical doctors and nurses. This bothered me more than her telling me about her participation in the swinging. It felt too personally intrusive, too close to the realm of possibility. It made me feel further isolated and somehow judged negatively by the patient, as if I were the outsider, not one of the group. Just as important, however, I felt relief. A “secret” had been revealed in the therapy that I found very frustrating. I now had renewed enthusiasm about the therapy and, with it, some small sense of control and guarded optimism that the therapy would now progress. I was concerned that this disclosure occurred 7 months into therapy and felt somewhat staged, but given her expressed desire for intimacy with women, this seemed to make sense at the time in terms of her fear that we would lose our connection.

Dr. Gabbard

The theme of voyeur/exhibitionist pairings heralded by the initial dream was given further elaboration after 7 months of therapy, when Ms. A described her practice of “swinging.” Although she was naive about this practice, Dr. Bennett strove to listen in a nonjudgmental and accepting way. She was a well-trained dynamic therapist who knew that at the heart of the psychodynamic approach was a profound respect for the patient’s individual right to live as she wished without imposing judgment or exhortations to change. After all, she reasoned, these are consenting adults, and there was no apparent harm to anyone. Her husband was not forcing Ms. A to engage in these foursomes. She enjoyed the sex and looked forward to it. Dr. Bennett was an observer of a fascinating narrative that was alien to her, lending a “cross-cultural” perspective to the therapy. She was a psychiatric version of a cultural anthropologist in her consulting room, eavesdropping on the sexual practices of creatures from a stratum of society that were not within her experience. When Ms. A noted that doctors and nurses were involved, however, Dr. Bennett started feeling uneasy, sensing that the stories were becoming a bit too close, encroaching on the borders of her own turf at an academic medical center. She was also drawn to the theatrical aspects of the presentation. It was like watching a television movie. The “staged” feeling of the proceedings helped her keep her distance. The pleasure in watching and showing, of course, can be understood as an enactment of the patient’s sexual inner world. As noted, the narrative of words and events has its nonverbal counterpart in the enactment of a sexualized scenario within the therapy. A fundamental psychodynamic principle is that each patient re-creates his or her internal object world in the therapeutic setting.

Dr. Bennett

Over time, Ms. A continued to reveal more detailed and explicit accounts of her sexual behavior. She and her husband often attended private swinging parties and frequented swinging clubs most weekends. Her description of these events featured rooms full of naked people drinking, talking, and involved in various sexual acts, some behind closed doors and many not. Threesomes and foursomes with spectators and participants changing roles were common. Ms. A actively sought out the largest, most overtly aggressive men at these events and would engage in intercourse, oral and/or anal sex with them, most often while her husband watched. The female-female sexual behavior also progressed to oral sex, with Ms. A always the “giver” (because she refused to be a recipient). This contrasted with her sex life with her husband, which never included oral sex. She continued to present this material in a dramatic manner that seemed designed to maximize its shock value. Ms. A recounted these events with a remarkably theatrical flair, including dramatic (dare I say “pregnant”?) pauses for effect when she appeared to be gauging my reaction. I did my best to respond in a consistent and nonjudgmental manner, but I worried increasingly about her safety. When I brought up this concern, she would rapidly reassure me that she always used condoms and that her husband was always there.

The sessions became even more graphic. She described “sex-a-thons” that involved going to conventional bars with her husband to pick up large, athletic men and bring them home. She and the chosen man would then have loud, aggressive sex for hours at a time, as she described it, while her husband listened downstairs. She also boasted repeatedly about her talent for pleasuring women in ways that amazed both the women themselves and their husbands. Often, she said, husbands or male partners would ask her afterward to teach them her technique. As I listened with fascination, I realized that the sessions had begun to feel like peep shows. I found myself actively disliking Ms. A for the first time and judging her somewhat contemptuously for participating in these sordid events. I also felt a degree of self-loathing for my role as a captive audience that had been dragged into these “show-and-tell” sessions. Nevertheless, I attempted to maintain the superficial demeanor of a concerned and interested neutral therapist. As her participation in this sexual lifestyle intensified, I worried that my nonjudgmental manner was actually encourag-
Dr. Gabbard

Dr. Bennett found herself in a real dilemma here. She was starting to feel dismayed by what she heard and contemptuous toward her patient. Yet she felt the therapy demanded her to be accepting of alternative sexual practices that were far from her experience and, from her perspective, morally questionable. Was her effort to present herself as accepting and nonjudgmental a form of collusion in which she is playing voyeur to the patient’s exhibitionist? Was she tactfully endorsing the behavior by not questioning it more vigorously? Yet if she did challenge it, Ms. A could, with some justification, accuse her of moralizing about her preferred sexual practices. After all, no children were involved, and no crime was being committed. Ms. A might simply clam up and stop talking about what she did in the privacy of her many “bedrooms.” Moreover, in response to her feelings of anger and contempt, at least to herself and to us. A psychodynamic therapist has no obligation toward political correctness in the domain of her private thoughts. She reacted in whatever way she felt and noted her feelings to herself in a way that may be useful in therapy. Was she reacting like others react to the patient? Was her reaction in some way idiosyncratic? As Sandler (1) suggested in his classic article on role-responsiveness, the therapist must maintain both free-floating attention and free-floating responsiveness. The decision to share those feelings with the patient was far more complicated and must be carefully considered in each individual case.

Dr. Bennett

Ms. A came in one day and casually mentioned that her husband had arranged a date for the two of them. She nonchalantly joked that her husband was like a pimp and chuckled about her analogy. When asked more about this, she casually stated that this was nothing new; her mother had done the same thing. She then described how her mother used to dress her up “like a doll” in a “princess dress” and take her to see her father at work at a bar. She would then be placed on a bar stool (she was too small to climb up herself) and was expected to smile, chat, and generally entertain the “regulars.” When her father wasn’t busy, she would be available to spend a couple minutes with him. Most frequently, however, she remembered these nights as consisting of older men of varying degrees of intoxication fawning over her, buying her pretty drinks, and telling her that she was a beautiful little girl. This went on throughout her childhood, starting from around the time she was 3 years old, at least once a week until close to adulthood. As an adolescent, her mother would watch from nearby. At this point, my face must have revealed some degree of amazement or concern because she paused and then added that she told her mother everything, even about the swinging. When I asked how her mother had responded, she said that her mother had told her about a lesbianencounterthat she had “almost” had but “didn’t go through with.” It seemed a source of pride to both mother and daughter that Ms. A had actually gone through with these experiences. I asked her how she felt telling her mother about such things. She snickered, stating that her mother had been dressing her in sexually provocative clothing since she was a young child. I then made a comment about how difficult it must be to be someone else’s plaything, and she shrugged, stating that she was used to it by now. In this regard, it is important to stress that Ms. A always expressed thorough enjoyment of the swinging, despite her husband having been the one to initiate these experiences.

Dr. Gabbard

Dr. Bennett could no longer contain her forced neutrality at this point. Ms. A noticed that she was amazed and concerned and took time out from her narrative to explain that she told her mother everything, as if to say to Dr. Bennett that she had to do the same with her therapist. In making the assumption that at some level Ms. A must have felt used and exploited by her husband in the present and by her mother in the past, Dr. Bennett made an empathic comment that it must have been difficult to be someone else’s plaything. Ms. A shrugged off the comment by reassuring her therapist that she was simply used to it, implying that it no longer bothered her.

The therapist was faced with a dilemma of increasing complexity here. Embedded in the notion of two consenting adults is freedom of choice. Two adult sexual partners are constitutionally free to do whatever they like when they are in private, so we mental health professionals should suspend judgment and let our patients seek whatever form of sexual gratification they like. But here’s the rub: is choice always free? Is a woman who is beaten during incestuous sexual relations with her father throughout her childhood then “free” to “choose” that lifestyle as an adult with abusive male partners? With our growing knowledge of the compelling need to repeat traumatic relationships (2, 3), can we really stand by in such cases and say that the patient is simply “choosing”?

Ms. A was clearly repeating a pattern from her childhood that felt obligatory at some level. To make her mother happy, she had dressed up like a party doll and functioned as a plaything to a host of inebriated men in a bar. Now she was reenacting that childhood scenario with her husband and to some extent with her therapist. When Dr. Bennett attempted to empathize with the degrading aspects of the experience, Ms. A brushed off her help. Sexually impulsive or compulsive patients are often treated today with an approach that emphasizes impulse-control training, 12-step groups, relapse prevention, cognitive restructuring, and social learning (4). This approach requires a patient who is willing to view her sexual behavior
as a problem and is interested in forming a collaborative
alliance with a therapist as part of a systematic treatment
plan. Ms. A expressed no interest whatsoever in changing
her behavior. What was Dr. Bennett to do?

Dr. Bennett

As Ms. A continued to relay her sexual experiences, a
degree of ambiguity gradually surfaced. During one ses-
sion over a year into the therapy, she stated that her hus-
band had complained that she always “zonked out” be-
fore a party or before going to a club. He was annoyed
because it appeared that she didn’t want to go out. Ms.
A then described a period of several hours’ duration that
typically occurred preceding swinging events. She would
essentially dissociate at these times, becoming nonre-
sponsive to others, focusing internally, and attempting
to self-soothe. She stated that she needed to do this to
“prepare” for the evening and that once she was at the
event, she was able to smile, socialize, and have sex. She
spontaneously connected this to a memory from her
childhood. Throughout her childhood and adolescence,
Ms. A was “strongly encouraged” by her mother to
model. As a young child, she was a runway model for a
department store chain. She vividly recounted an experi-
ence when she was 3 years old: it was before a fashion
show, and she was backstage with her mother. She was
crying and screaming that she did not want to do it. She
had tears running down her face, and her makeup and
hair were getting ruined. Her mother became very angry,
yelling and swearing at her to “do it.” She struggled with
her mother, still crying that she didn’t want to do it. Her
mother pushed her onto the runway, yelling in a more
hushed tone, and suddenly she was out on the runway,
and everyone was staring at her. She paused, remember-
ing the dress as pink and black with ruffles, and then
stopped. I asked what happened next. She sarcastically
responded, “What do you think happened? I stopped
crying and did it.” She then casually added that it was
the same with swinging, that she needs to prepare for
the performance, but once “on,” she can “do it.”

Ms. A often relayed snippets of her childhood that
were as dramatic and “shocking” in both delivery and
content as those regarding her sexual exploits. Often-
times these memories were as uncomfortably tantalizing
as listening to her describe her sexual exploits. Fre-
cently, they seemed intertwined. Whereas I often vis-
ited my internal struggle over my attempt to remain
nonjudgmental regarding her swinging, I had no hesi-
tancy whatsoever voicing my judgment regarding her
childhood experiences with her mother. I could not help
but think of my own daughter when listening to Ms. A at
these times and knew in my gut that my patient’s child-
hood experiences were unequivocally wrong. When Ms.
A relayed such childhood memories, I responded empa-
thetically and supportively, emphasizing how horrible
such things must have been for her and validating her
view of her mother as a very controlling, selfish, and, at
times, downright sadistic woman. At first, Ms. A resisted
this intervention, stating that she loved her mother and
knew she had done “the best she could.” I did not dis-
pute this but kept emphasizing how horrible the scenar-
ios were that she described and how hard it must have
been to develop one’s own self when one’s primary care-
giver had such a dominant personality. This approach
seemed to calm her, and she seemed to consider what I
was saying for the first time. In terms of content, she
gradually shifted to talking less about her sexual exploits
and more about her childhood. She became more open
about her insecurities regarding being female, including
her discomfort with her body and its femininity.

Dr. Gabbard

Here Dr. Bennett found an ingenious way to form a ther-
apeutic alliance with Ms. A. Rather than taking a judgmen-
tal stand on Ms. A’s current sexual practices, she focused
her attention on childhood experiences that we can view
as the antecedents of her current sexual behavior. Here
she could join the patient in a sense of outrage about a
“stage mother” who forced her own agenda on her daugh-
ter. By frankly sharing her horror at the modeling stories,
Dr. Bennett facilitated a feeling of validation in Ms. A. This
sense of having her experience affirmed helped build trust
and a sense of feeling understood in the therapy. While it is
often risky to take sides against a patient’s mother, Dr.
Bennett finessed this problem by focusing on the horror of
the experiences rather than any sort of absolute “evil” in
the patient’s mother.

Dr. Bennett

In terms of the transference, the room felt warmer.
The countertransference thawed a bit, too. I began to
like Ms. A. I worried about her over the weekends and
wanted to keep her safe and protected. To a large extent,
I felt like I was mothering an adolescent. She continued
to discuss material that was just as fantastic and sexually
explicit as ever, but her motive seemed to be less to titil-
late or “shock” and more to share. She appeared to hear
me when I expressed my dismay or concern. Finally, her
behavior started to change, and her general impulsivity
(angry outbursts, reckless driving, and alcohol use) be-
egan to diminish.

Ms. A had a difficult time with my canceling appoint-
ments, regardless of the advance notice. Despite my con-
sistently outlining coverage in my absence and/or giving
her a cell phone number to call, she developed a pattern
of canceling her last one or two appointments before my
scheduled absences. During several sessions before one
such scheduled absence, we delineated a plan for a
weekend during which her parents, her husband, and I
would all be gone. Ms. A had stated repeatedly that she
knew that she would just want to “go out...act crazy...do
something impulsive.” She described herself as “horny”
and did not think she could resist the urge to have sex.
In spite of these statements, I sensed an ambivalence to
act. Her tone was less passionate, her words lacked cer-
tainty, and she seemed to be looking for a way out. It felt
like she was taunting me in an adolescent manner, as if
daring me to discipline her. We discussed alternative be-
aviors that she enjoyed, such as reading and playing
with her pets. I asked about masturbation with some
reservation, and she snickered, stating that she had al-
ready thought of that and had plenty of “toys” to keep
her occupied, adding, “but it just isn’t the same.” I em-
phasized that it was a lot safer and positively reinforced
how well she was doing caring for herself, thinking and
talking rather than doing. I also asked her to picture me
telling her to be safe and to tell herself “don’t do it” if
she thought she was going to do something impulsive
and potentially self-harmful. She laughed and said she
would. She then canceled the last session before this
weekend, stating in her voice mail message that she
knew I would think she was canceling because I was go-
ing away but really she had other things she needed to
do and no other time to do them. At our first session af-
after the weekend, she laughed when she described what
she had done that weekend—stayed home alone, played
with the pet, and read. She stated that she had come
close to going out and being “wild” several times but was
overly proud that she hadn’t. I was proud of her and de-

erived pleasure out of being her “good mommy.”

More than 2 years into the therapy, Ms. A stopped
swinging. This correlated with her attempts to get preg-
nant. She originally did not talk of giving up this lifestyle
tirely or permanently but, rather, temporarily to
“keep healthy” while she became pregnant. On several
occasions, I asked her how she planned to incorporate
parenting into this lifestyle. She stated matter-of-facty
that many of the swingers had kids and that some had
even been present (in different parts of the house) at pri-

date parties. Although I never verbalized any disapproval
regarding this, I am sure that my nonverbal distress was
clearly apparent to this very sensitive patient. Over time,
she talked more about the kind of parent she wanted to
be and her concern regarding her husband as a parent.
She maintained her confidence that she would be a good
mother and ultimately decided she would not want to
do both. Concurrently, she and her husband moved from
an apartment to a new condominium. Ms. A would talk
about her new home, the tree-lined sidewalks, the quiet
street, and how a baby was all that was missing from this
“perfect” life. This appeared to be a kind of substitute
fantasy. She defined this time period as “calm” but al-
ways voiced distrust of it, unsure whether or when “the
other shoe [would] drop.” Outwardly, she was containing
her impulsivity and becoming more productive. She
completed coursework and passed the certification ex-
amination to become a personal trainer and started
work part-time at a women’s fitness center. She found
working exclusively with women very rewarding and
genuinely enjoyed the work. Inwardly, she still “craved”
sex and missed the excitement and physical pleasures of
the swinging. She talked about this, likening herself to an
“addict,” stating that she was learning to resist her
“urges” but that sometimes it was “1 hour at a time.”

Her husband was very unhappy with her decision to
give up swinging, despite her unwavering and even in-
creased desire to have sex with him. He continued to
pressure her to start swinging again and stated that he
no longer wanted children. She acquiesced to his de-
mands for no children “for the sake of the marriage” but
did not start swinging again. It had been almost 1 year
since she last attended or participated in a swinging
event. She and her husband had very little sex and main-
tained a civil roommate-like day-to-day existence. De-
spite her desire for more sex and her stated desire to
save the marriage, she appeared unwilling to reenter the
swinging lifestyle.

She then purchased a vibrator. She actually blushed
when telling me this and quickly noted the irony, that of
all of the things she had told me, this was the first thing
that had embarrassed her. Unlike previous sex toys, this
vibrator was purchased exclusively for herself, and she
made the decision without consulting or informing any-
one. She planned on using it by herself, not as a prop
with her husband or others. I again emphasized how
well I thought she was doing and how far she had come
in terms of containing her impulsivity and keeping her-
self safe. I then told her that she should pat herself on
the back, to which she responded, “No. You should pat
yourself on the back,” and we both, literally, patted our-
selves on our backs.

Such exchanges started out somewhat in jest when Ms.
A’s behavior became more and more self-contained and
have become more frequent as she has improved: I con-
gratulate her for her good work. She typically responds
that it was my doing. We both agree that it was our good
work, and we both pat ourselves on our backs. This time,
however, she finished her patting and told me that at the
end of each exercise class that she teaches, she tells the
women to pat themselves on the back, and at the end of
her day (seven classes some days), she pats herself on the
back as well. I felt very touched by this image, not just in
terms of professional pride for the work we have done
but on a more visceral level, like when I watch my
daughter doing or saying something that seems all too
familiar and then realize it is because she is, on a very
concrete level, emulating me.

Dr. Gabbard

In this extraordinary piece of therapeutic work, Dr. Ben-
nett had to avoid a particular minefield involving the
repetition of specific patterns of the patient’s internal ob-
ject relations in the therapeutic relationship. Ms. A was
haunted by a sadistic, demanding object representation
that insisted on compliance with its wishes. While the
original model for this internal object may have been her
mother, it was reenacted on a regular basis with her hus-
bond in that role. Her corresponding self-representation
was one of a compliant child who would deny her rage
and helplessness and “prepare” herself for submission to the
demands. The therapist’s minefield is the almost magnetic
pull to start telling the patient what to do and what not to
do, which feeds right into the familiar paradigm in the pa-
tient’s internal object world. To some extent, this is un-
avoidable, but to her credit, Dr. Bennett largely avoided
the position of criticizing Ms. A for her sexual practices by
focusing primarily on her personal concern about Ms. A’s
safety. Although the patient obviously read expressions of
concern and disapproval in Dr. Bennett’s face, the ther-
papist did not assume the role of a critical, nagging mother
who told her that what she was doing was wrong. Hence,
Ms. A began to see her as a “good mommy” who had her
best interests in mind, a new object-relations paradigm
that was foreign to Ms. A. One way to understand the
mode of therapeutic action in this case was to assume that
a new neural network, consisting of a caring, concerned
object representation and a self-representation that en-
gaged in self-care and gained approval from doing so, was
gradually strengthened as the old neural networks of
problematic relatedness were relatively weakened. An-
other way to conceptualize what happened is that Ms. A’s
behavior was gradually transformed from egosyntonic to egodystonic. In other words, what she initially insisted was entirely acceptable and fun for her was converted to problematic status in her mind. We can speculate that Ms. A projected one-half of her ambivalence into Dr. Bennett—namely, the part that regarded the sexual practices negatively. Over time, Dr. Bennett contained that portion of Ms. A’s feelings until Ms. A could take that back and “own” it as her own feelings. While both transference and countertransference were important constructs to understand what was happening, Dr. Bennett did not actively interpret transference to Ms. A. She also encouraged a sense of agency in the patient by refusing to take credit for all the improvements that had taken place, thus assisting the patient in developing a sense of self separate from being a passive recipient of the wishes of others.

References
William Blake (1757–1827)

William Blake was both a great poet and a great painter. Blake invented a method of relief engraving that was able to combine his poetry and paintings so that they formed a unified work. In the West, only Michelangelo had previously demonstrated such singular talent in both painting and poetry, but unlike Blake he did not combine them. Blake was one of the progenitors of the revolutionary Romantic movement, and his verse was admired by Coleridge and Wordsworth.

Stanzas by Blake such as the following have become an imperishable part of the Western canon:

To see a World in a Grain of Sand
And a Heaven in a Wild Flower
Hold Infinity in the palm of your hand
And Eternity in an hour
—from “Auguries of Innocence”

Long before Heinz Kohut and self psychology, Blake captured the essential tragedy of narcissistic pathology:

Love seeketh only Self to please,
To bind another to its delight,
Joys in another’s loss of ease,
And builds a Hell in Heaven’s despite
—from “The Clod and the Pebble”

Blake’s father owned a hosiery and haberdashery shop in London. As was common for tradesmen, the family lived above the shop. Blake was the third child of six. He was raised in the Protestant Dissenting tradition, one that emphasized private devotion and individual conscience separate from the authority of priest and church. In the lexicon of the 18th century, the term for extreme dissent was “enthusiasm” (possessed by God). Later in life, Blake, who was deeply religious, referred to himself as an “enthusiastic hope-fostered visionary.” He demonstrated artistic and literary talent from an early age, and his enlightened parents enrolled him in a drawing school at the age of 10. He was educated as a commercial engraver and was later admitted as a student to the prestigious Royal Academy. In 1782, Blake married a woman from a working-class family. This was a childless but mutually devoted marriage that lasted 45 years. Blake’s artistic career underwent many financial vicissitudes, and though he had a number of aristocratic patrons who recognized his originality, his true genius was not fully appreciated until after his death. He may have died from liver failure secondary to biliary cirrhosis induced by chronic copper ingestion during his etching copper plates for his engravings.

From his childhood onward Blake saw visions. His biographer, Bentley, recounts, “Once his mother beat him for running in and saying that he saw the prophet Ezekial under a Tree in the Field” (1, p. 19) and, “Later when he was eight or ten, one day as he was walking—he saw a tree filled with angels, bright angelic wings be-spangling every bough like stars” (p. 19). His wife commented in 1810, “I have very little of Mr. Blake’s company; he is always in Paradise.” Blake’s recurrent hallucinatory visions suffused his poetry and painting, especially his great epics and their accompanying watercolors and engravings.

Alongside his ecstatic visions, Blake was prone to fits of severe depression. In 1800, he recounted a descent into “a Deep pit of Melancholy, Melancholy without any real reason for it.” These episodes were often followed by periods of “illumination” and intense creativity. This is highly suggestive of bipolar illness, albeit a mild form that did not disrupt his enormous creative achievement and may have been central to his transcendent artistic vision.

Reference


PETER J. BUCKLEY, M.D.

Address correspondence and reprint requests to Dr. Buckley, 336 Central Park West #5A, New York, NY 10025; pbuckley@montefiore.org (e-mail).

Image courtesy of the Tate Gallery, London/Art Resource, New York.
Borderline Personality Disorder in Clinical Practice

Carolyn Zittel Conklin, Ph.D.
Drew Westen, Ph.D.

Objective: Most studies of borderline personality disorder have drawn patients from among hospital inpatients or outpatients. The aims of this study were to examine the nature of borderline personality disorder patients in everyday clinical practice and to use data from a sample of borderline personality disorder patients seen in the community to refine the borderline construct.

Method: A random national sample of 117 experienced psychiatrists and psychologists from the membership registers of the American Psychiatric Association and American Psychological Association provided data on a randomly selected patient with borderline personality disorder (N=90) or dysthymic disorder (N=27) from their practice. The clinicians provided data on axis I comorbidity, axis II comorbidity, and adaptive functioning, as well as a personality description of the patient using the Shedler-Westen Assessment Procedure-200 (SWAP-200) Q-sort, an instrument designed for assessment and taxonomic purposes. Analyses compared borderline personality disorder and dysthymic disorder groups on variables of interest and aggregated SWAP-200 items across all borderline personality disorder patients to create a composite portrait of borderline personality disorder as seen in the community.

Results: The borderline personality disorder sample strongly resembled previously studied borderline personality disorder samples with regard to comorbidity and adaptive functioning. However, the SWAP-200 painted a portrait of borderline personality disorder patients as having more distress and emotion dysregulation, compared to the DSM-IV description.

Conclusions: Borderline personality disorder patients in research samples are highly similar to those seen in a cross-section of clinical practice. However, several studies have now replicated a portrait of borderline personality disorder symptoms that places greater weight than the DSM-IV description on the intense psychological pain of these patients and suggests candidate diagnostic criteria for DSM-V.

Since the first research using a standardized interview for borderline personality disorder patients two decades ago (1), an immense body of research has emerged on the nature and etiology of borderline personality disorder. Most studies have drawn subjects from groups of outpatients or inpatients, usually associated with academic training departments (e.g., references 2–19). To what extent these patients, who are likely to have symptoms on the more disturbed end of the borderline spectrum, resemble the range of borderline personality disorder patients seen in everyday practice is largely unknown.

The aims of the current study were twofold. The first was to describe the nature of borderline pathology seen in clinical practice. We compared data from prior studies with data from a random national sample of borderline personality disorder patients treated in the community on three sets of criteria: axis I comorbidity, axis II comorbidity, and adaptive functioning. Gunderson’s review (20) indicated that the axis I disorders most frequently found in borderline personality disorder patient samples are dysthymic disorder, major depression, substance abuse, posttraumatic stress disorder, and eating disorders and that at least one-half of borderline personality disorder patients have major depressive disorder, dysthymia, or both. Although borderline personality disorder has been found to have high rates of comorbidity with virtually all axis II disorders, the highest diagnostic overlap appears to be with histrionic and avoidant personality disorders (20, 21). With regard to adaptive functioning, research findings have associated borderline personality disorder with self-injurious behavior such as skin cutting and burning and with psychiatric hospitalizations, suicidality, difficulty maintaining relationships, and difficulty maintaining appropriate employment. We thus expected to see similar patterns of findings in a community clinical sample if the descriptions of borderline personality disorder generated from hospital inpatients and outpatients generalize.

The second aim was to describe the personality characteristics of borderline personality disorder patients by using a large, relatively comprehensive item set and to refine the borderline construct empirically by using a broad sample of borderline personality disorder patients seen in the community. In a prior study (22, 23), a large random national sample of experienced clinicians described a
personality disorder patient by using the Shedler-Westen Assessment Procedure-200 (SWAP-200) (22), a clinician-report personality pathology Q-sort instrument that includes items reflecting the roughly 80 DSM-IV criteria for all current axis II diagnoses as well as 120 additional items that provide candidate criteria for refining current diagnoses (i.e., potential alternative diagnostic criteria). Of 530 clinician-participants, 43 described a patient with borderline personality disorder. Among the items most characteristic of the borderline personality disorder patients in this sample were several that mirrored DSM-IV criteria. Other items, however, appeared to be more characteristic of the average borderline personality disorder patient than several of the DSM-IV criteria, notably items describing intense and poorly modulated affect and profound dysphoric affect. The data suggested that intense dysphoric affect is a core, rather than co-occurring, feature of borderline personality disorder. Similar findings emerged in a prior study that used the SWAP-167, the progenitor to the SWAP-200 (24). The results of these studies were in keeping with Gunderson's finding that chronic major depression and chronic feelings of helplessness, hopelessness, worthlessness, guilt, loneliness, and emptiness appear to be central to the disorder (20).

In the present study, we asked a random national sample of experienced clinicians to provide data on phenomenology, comorbidity, and adaptive functioning in a randomly selected patient with DSM-IV-diagnosed borderline personality disorder, and we used data from the SWAP-200 Q-sort to develop an empirical portrait of the personality functioning of the average patient with borderline personality disorder. Given that the instrument includes items assessing all current axis II criteria, if personality descriptors that were not among the DSM-IV diagnostic criteria appeared to be more diagnostic than the current criteria in a sample specifically selected for meeting those criteria, and if these descriptors replicated those found to be more descriptive of borderline personality disorder patients in prior research, these findings would suggest the need for refining the borderline construct to mirror more closely the nature of patients seen in the community.

Method

The present investigation relied on practice network methods to address taxonomic and other basic science questions. Elsewhere we addressed in detail the rationale for this clinician-report method, including its advantages and limitations (see references 22, 25–30). In brief, clinicians are experienced observers who observe patients longitudinally and in depth. Although unstructured clinical judgments have been shown to have poor reliability and validity, a host of recent studies suggested that clinicians can provide highly reliable and valid data when they quantify their judgments using psychometric instruments and that their data predict clinical outcomes better than their beliefs about or theories of psychopathology (see, e.g., references 24, 34).

Participants and Procedures

Participants were 117 clinicians who constituted a random national sample of experienced psychiatrists and psychologists from the membership registers of the American Psychiatric Association and the American Psychological Association. Initial letters to clinicians described the study, presented them with the DSM-IV diagnostic criteria for borderline personality disorder and dysthymic disorder (which was selected as a comparison condition), and asked them to complete a postcard indicating whether they had at least one borderline personality disorder or dysthymic disorder patient in their practice who met the inclusion and exclusion criteria. Based on their postcard responses, we assigned the clinicians to describe either a borderline personality disorder patient (N=90) or a dysthymic disorder patient (N=27), again presenting them with the DSM-IV diagnostic criteria for borderline personality disorder and/or dysthymic disorder to ensure close attention to the diagnostic criteria. To ensure random selection of patients, we asked the clinicians who reported having more than one appropriate patient to consult their calendars and select the patient they saw most recently who met the study criteria. For the dysthymic disorder group, we asked clinicians to describe a current patient who met the DSM-IV criteria for dysthymic disorder and who had no diagnosable DSM-IV personality disorder and no more than three DSM-IV diagnostic criteria for borderline personality disorder. For patients in both groups, we asked clinicians to select a female patient (to avoid the confounding factor of gender and to maximize power, because 75%–80% of patients who receive a diagnosis of borderline personality disorder are female [20, DSM-IV]) between ages 18 and 55 years (to avoid the confounding factors associated with adolescent and late-life personality disorder diagnosis) whom they had seen for a minimum of eight sessions and a maximum of 2 years (to guarantee that they knew the patient well while minimizing the likelihood of substantial personality change in treatment) and who did not have a psychotic disorder. We asked clinicians to select a current psychotherapy patient to maximize the likelihood of their being able to provide detailed personality assessments. (We selected a comparison group of patients with dysthymic disorder because patients with depression have been the most common comparison group in studies of personality disorders, and patients with dysthymic disorder have enduring, moderate depression that is also common in patients with borderline personality disorder.) To maximize participation, we gave clinicians the option to participate by pen and paper or on our interactive web site (http://www.psychsystems.net). Consistent with the literature on computerized versus paper administration of questionnaires (35), we found no systematic differences between responses with the two methods.

Before analyzing the data, we excluded data on patients who were extreme outliers in age or length in treatment beyond the parameters we requested, data on patients who did not meet the diagnostic criteria for the borderline personality disorder or dysthymic disorder groups, and data suggesting extreme carelessness in responding (e.g., multiple pages not completed). To maximize power, however, we retained patients who exceeded within reasonable bounds the maximum limit for age (two patients whose ages were in the range of 55–61 years) and time in treatment (six patients whose time in treatment ranged from 25 to 48 months). Further, because several dysthymic disorder patients were one criterion short of the diagnostic criteria for the disorder or met the criteria for multiple personality disorders, we were faced with decisions about the “purity” of the dysthymic disorder sample. We ultimately chose to retain patients who had chronic depression if they were within one criterion of the dysthymic disorder diagnosis and to retain dysthymic disorder patients who met the DSM-IV criteria for a non-borderline-personality-disorder diagnosis (mostly avoidant and schizoid personality disorders) to maximize the number of subjects and the generalizability of the
sample. The decision to include non-borderline-personality-disorder patients actually rendered findings more conservative and increased external validity, given the high rates of comorbidity (60%) for dysthymic disorder and personality disorders in prior research (36, 37). (In fact, we reran all analyses without the eight patients who did not meet the criteria, and significance values improved in three cases and decreased from 0.01 to 0.05 in one. However, to preserve consistency with other reports of data from this sample, we chose to avoid excluding these subjects for some analyses but not for others.)

**Measures**

Clinicians completed the following measures, presented in the following order. (We included other instruments for other studies but do not describe them here.)

**Clinical Data Form.** The Clinical Data Form was used to assess a range of variables relevant to demographics, diagnosis, adaptive functioning, developmental history, and family history of psychopathology. This measure was developed over several years and used in a number of studies (see reference 38). The sections of the Clinical Data Form that were relevant to this study ask clinicians to provide basic demographic data on themselves and the patient, as well as information pertaining to the patient’s diagnosis and adaptive functioning. Prior research found such ratings to correlate strongly with ratings made by independent interviewers (28, 33, 39).

**SWAP-200.** The SWAP-200 is a 200-item Q-sort designed to assess personality and personality pathology (e.g., references 22, 24, 27, 38). (A Q-sort is a set of statements printed on separate index cards, in this case, statements about personality and personality dysfunction.) An experienced clinical observer sorts the cards into eight piles, thereby assigning each of the 200 descriptive items a numerical score ranging from 0 (for items least descriptive of the patient) to 7 (items most descriptive of the patient). Items for the SWAP-200 were derived from a number of sources, including DSM-III-R and DSM-IV axis II criteria, clinical literature on personality disorders, research on normal personality traits and psychological health, pilot interviews, and the feedback of more than 1,000 clinicians.

Development of the item set was an iterative process that followed standard psychometric methods, such as eliminating redundant items, items with minimal variance, and so forth. The Q-sort items provide a standardized clinical language that allows for clinicians’ assessments to be quantified, compared with those of other clinicians, and analyzed statistically.

Research thus far has supported the validity and reliability of the SWAP-200 in predicting numerous external criteria, such as suicide attempts and history of psychiatric hospitalizations, adaptive functioning assessed by measures such as the Global Assessment of Functioning Scale (GAF) from the DSM-IV, diagnoses based on interviews, and developmental and family history variables (e.g., references 25, 27, 34). The SWAP-200 has been used for taxonomic purposes in multiple studies (e.g., for empirically deriving personality diagnoses from large samples of adult and adolescent patients [references 27, 40]).

**Axis I checklist.** Clinicians completed a present/absent checklist of the most common axis I DSM-IV diagnoses in borderline personality disorder reported in the literature, including major depressive disorder, dysthymic disorder, bipolar disorder, generalized anxiety disorder, panic disorder, anorexia nervosa (restricting type), anorexia nervosa (binge-eating/purging type), bulimia nervosa (purging type), bulimia nervosa (nonpurging type), alcohol abuse/dependence, prescription drug abuse/dependence, and illicit drug abuse/dependence. For a subset of disorders in which we were particularly interested (dissociative disorders, posttraumatic stress disorder, and dysthmic disorder), we asked clinicians to make present/absent distinctions on each DSM criterion for each disorder, which allowed us to apply DSM-IV algorithms to identify patients who met the diagnostic criteria.

**Axis II checklist.** Clinicians completed a checklist containing the criteria for the DSM-IV personality disorders, randomly ordered, so that we could assess axis II pathology both dimensionally (number of symptoms endorsed) and categorically (applying DSM-IV cutoffs), again without relying on clinicians’ global diagnoses. Similar methods have been employed by other researchers, such as Blais and Norman (41).

**Data Analysis**

The aims of the study were 1) to examine the nature of borderline personality disorder patients in everyday clinical practice and their resemblance to borderline personality disorder patients seen in research studies and 2) to see if we could identify candidate diagnostic criteria for a revision of DSM-IV. Thus, after conducting diagnostic validity checks, we examined their similarity to borderline personality disorder patients prototypically described in research accounts in terms of axis I and axis II comorbidity and adaptive functioning. We hypothesized that the borderline personality disorder group, compared to the dysthmic disorder group, would demonstrate lower levels of various indices of adaptive functioning in chi-square analyses for categorical variables and t tests for dimensional variables. We report Pearson’s r as an effect size estimate throughout. (For interpretations of r as a measure of effect size, see references 42, 43.) To create a composite personality portrait of borderline personality disorder in everyday practice and to see if we could identify candidate criteria for the disorder that might be more identifying than the DSM-IV criteria, we aggregated SWAP-200 item scores across all borderline personality disorder patients.

**Results**

**Demographics**

Of the 117 clinician-participants, 19% (N=22) were psychiatrists and 81% (N=95) were psychologists (the latter responded at a much higher rate to the initial solicitation); 42% (N=49) were female. The majority worked at least part time in private practice (88%, N=103), although many worked in other settings as well, with 18% (N=21) working in a clinic; 26% (N=30) in outpatient, inpatient, or partial hospital settings; 8% (N=9) in a forensic setting; and 16% (N=19) in other settings. Clinicians were diverse in theoretical orientation, with 21% (N=24) describing their psychotherapeutic orientation as cognitive behavioral or behavioral, 44% (N=52) as psychodynamic or psychoanalytic, 32% (N=37) as eclectic, and 3% (N=4) as other.

Patients were an average age of 38 years (SD=10.14). The sample was predominantly Caucasian (88%, N=103); about 5% (N=6) were Hispanic, and the remainder were African American, Asian, or another ethnicity. Patients were primarily working class (33%, N=39) and middle class (47%, N=55), with educational attainment ranging from high school degree (15%, N=18) to having some graduate education (26%, N=30). Patients had been in treatment for an average of 12 months (SD=7.8), so the clinicians knew them well.
TABLE 1. Axis I Diagnoses in Patients With Borderline Personality Disorder and Dysthymic Disorder From the Practices of a Random National Sample of 117 Clinicians

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Patients With Borderline Personality Disorder (N=90)</th>
<th>Patients With Dysthymic Disorder (N=27)</th>
<th>Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Major mood disorders</td>
<td>85</td>
<td>94.4</td>
<td>27</td>
</tr>
<tr>
<td>Major depressive disorder</td>
<td>67</td>
<td>74.4</td>
<td>6</td>
</tr>
<tr>
<td>Dysmthetic disorder</td>
<td>29</td>
<td>32.2</td>
<td>27</td>
</tr>
<tr>
<td>Bipolar I disorder</td>
<td>16</td>
<td>17.8</td>
<td>0</td>
</tr>
<tr>
<td>Bipolar II disorder or cyclothymia</td>
<td>9</td>
<td>10.0</td>
<td>0</td>
</tr>
<tr>
<td>Antisocial disorders</td>
<td>40</td>
<td>66.7</td>
<td>7</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>42</td>
<td>46.7</td>
<td>4</td>
</tr>
<tr>
<td>Posttraumatic stress disorder</td>
<td>28</td>
<td>31.1</td>
<td>3</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>24</td>
<td>26.7</td>
<td>0</td>
</tr>
<tr>
<td>Dissociative disorders</td>
<td>48</td>
<td>53.3</td>
<td>0</td>
</tr>
<tr>
<td>Depersonalization disorder</td>
<td>40</td>
<td>44.4</td>
<td>0</td>
</tr>
<tr>
<td>Dissociative identity disorder</td>
<td>10</td>
<td>11.1</td>
<td>0</td>
</tr>
<tr>
<td>Dissociative amnesia</td>
<td>16</td>
<td>17.8</td>
<td>0</td>
</tr>
<tr>
<td>Substance abuse/dependence</td>
<td>32</td>
<td>35.6</td>
<td>2</td>
</tr>
<tr>
<td>Eating disorders</td>
<td>22</td>
<td>24.4</td>
<td>1</td>
</tr>
<tr>
<td>Anorexia nervosa</td>
<td>10</td>
<td>11.1</td>
<td>0</td>
</tr>
<tr>
<td>Bulimia nervosa</td>
<td>10</td>
<td>11.1</td>
<td>1</td>
</tr>
</tbody>
</table>

* Data were provided by experienced psychiatrists (N=22) and psychologists (N=95) identified from the membership registers of the American Psychiatric Association and the American Psychological Association. Each clinician randomly selected one patient with borderline personality disorder or dysthymic disorder from his or her practice. Only comorbid diagnoses with at least 5% prevalence in one or both groups are included. Effect size estimate (see references 42, 43 for interpretations of $r$ as a measure of effect size). *p<0.05. **p<0.01. ***p<0.001.

TABLE 2. Comorbid Axis II Diagnoses in Patients With Borderline Personality Disorder (N=90)

<table>
<thead>
<tr>
<th>Personality Disorder Diagnosis</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paraphrenia</td>
<td>33</td>
</tr>
<tr>
<td>Schizoid</td>
<td>22</td>
</tr>
<tr>
<td>Schizotypal</td>
<td>14</td>
</tr>
<tr>
<td>Antisocial</td>
<td>25</td>
</tr>
<tr>
<td>Histrionic</td>
<td>23</td>
</tr>
<tr>
<td>Narcissistic</td>
<td>18</td>
</tr>
<tr>
<td>Avoidant</td>
<td>48</td>
</tr>
<tr>
<td>Dependent</td>
<td>29</td>
</tr>
<tr>
<td>Obsessive-compulsive</td>
<td>14</td>
</tr>
</tbody>
</table>

Validity Check

As a validity check, we conducted two t tests to compare borderline personality disorder and dysthymic disorder patients on two measures of borderline personality disorder. Using the clinicians' 7-point ratings of the extent to which patients matched the borderline personality disorder construct, we found that clinicians rated borderline personality disorder patients significantly higher than dysthymic disorder patients (t=27.72, df=29.15, p<0.001, r=0.97). The same pattern emerged when we instead used SWAP-200 borderline personality disorder scale scores as the criterion variable (t=12.26, df=115, p<0.001, r=0.75) (17).

Borderline Personality Disorder as Seen in Everyday Practice

Comorbidity. Table 1 reports the frequency of comorbid axis I conditions in the two groups. The borderline personality disorder group distinguished itself both by the sheer number of comorbid diagnoses on average and by the specific diagnoses that have commonly been reported in reviews of studies of borderline personality disorder patients (20, 21).

We assessed axis II comorbidity by applying DSM-IV algorithms to the axis II symptom checklist data. As Table 2 shows, comorbidity was substantial, with the pattern once again strongly resembling that reported in prior studies (20, 21). (We do not report axis II comorbidity data for the dysthymic disorder patients because we requested that the clinicians provide data for dysthymic disorder patients without personality disorder disorders.)

Adaptive functioning. As Table 3 and Table 4 show, borderline personality disorder patients functioned significantly more poorly than dysthymic disorder patients on multiple indices of adaptive functioning. For example, on a 7-point rating of chronic level of personality functioning based loosely on Kernberg's model of levels of functioning (44) (using four anchors: "psychotic," "personality disorder," "substantial problems," and "high-functioning"), clinicians rated the dysthymic disorder patients more than 2 points (and three standard deviations) higher than the borderline personality disorder patients. The one exception was the number of confidants, a measure of social support, which makes sense in light of the association of borderline personality disorder with the personality trait of extroversion (45).

Most (70%, N=63) of the borderline personality disorder patients had attempted suicide. Attempters on average had made 3.89 attempts (SD=6.72), with the severity of the most dangerous attempt rated on average as "moderate, requiring medical attention." Most (63%, N=57) of the borderline personality disorder patients had at least one psychiatric hospital admission, and those who had a history...
of hospitalization had an average of 3.67 admissions (SD=3.92). More than one-half of the borderline personality disorder patients (52%, N=47) had self-injured. Of the 47 borderline personality disorder patients who self-injured, 81% (N=38) cut, 23% (N=11) burned, and 13% (N=6) severely scratched or tore their skin; an additional 26% (N=12) had repeated accidents.

Patients with borderline personality disorder showed generally poor relational functioning across several measures. Of particular interest, 41% (N=37) had been in abusive relationships in adulthood, with the majority in the role of victim (60%, N=22), a substantial minority in the roles of both victim and perpetrator (38%, N=14), and only one exclusively in the role of perpetrator. Nearly one-third (32%, N=29) had been the victim of rape or sexual assault in adulthood (see reference 46), and, for borderline personality disorder patients who reported any such incident, rape or sexual assault occurred on average 2.75 times (SD=3.66).

Identifying Candidate Diagnostic Criteria

To construct a composite personality portrait of borderline personality disorder in clinical practice, we aggregated scores for each of the 200 items of the SWAP-200 across all 90 borderline personality disorder patients by taking the mean across subjects and then arraying the items in descending order of magnitude (i.e., beginning with the items most characteristic of the borderline personality disorder patients). Table 5 lists the items with the highest average rankings. (Items ranked significantly higher for borderline personality disorder patients than for dysthymic disorder patients are starred in Table 5.)

Most DSM-IV diagnostic criteria for borderline personality disorder were represented among the SWAP-200 items with the highest average rankings. However, several SWAP-200 items with the highest average rankings were not reflected in the DSM-IV criteria. Table 6 presents the non-DSM-IV items that were rank-ordered among the top 20 items, all of which also ranked within the 20 most descriptive items in the largest prior sample to date in which this method was used (23), suggesting that the findings are robust and not attributable to particular characteristics of the present sample. In general, these items captured negative affect, emotion dysregulation, and poor self-esteem or self-loathing. Perhaps most striking, the two items most descriptive of borderline personality disorder patients in both samples were “Tends to feel unhappy, depressed, or despondent” and “Emotions tend to spiral out of control.”

Discussion

Primary Findings

Virtually all research on patients with borderline personality disorder has studied samples from hospital inpatient units or outpatient clinics. This study compared borderline personality disorder patients with dysthymic disorder patients treated in the community to see whether the phenomena observed in prior samples characterize borderline personality disorder patients as treated in the

### TABLE 3. Adaptive Functioning of Patients With Borderline Personality Disorder and Dysthymic Disorder

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients With Borderline Personality Disorder (N=90)</th>
<th>Patients With Dysthymic Disorder (N=27)</th>
<th>Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Lost job in past 5 years</td>
<td>46</td>
<td>51.1</td>
<td></td>
</tr>
<tr>
<td>History of suicide attempts</td>
<td>63</td>
<td>70.0</td>
<td></td>
</tr>
<tr>
<td>History of self-injury</td>
<td>47</td>
<td>52.2</td>
<td></td>
</tr>
<tr>
<td>Psychiatric hospitalization</td>
<td>57</td>
<td>63.3</td>
<td></td>
</tr>
<tr>
<td>Rape or sexual assault</td>
<td>29</td>
<td>32.2</td>
<td></td>
</tr>
<tr>
<td>Abusive romantic relationship</td>
<td>37</td>
<td>41.1</td>
<td></td>
</tr>
</tbody>
</table>

* Effect size estimate (see references 42, 43 for interpretations of r as a measure of effect size).
*p<0.05. **p<0.01. ***p<0.001.

### TABLE 4. Clinicians’ Ratings of Adaptive Functioning in Patients With Borderline Personality Disorder and Dysthymic Disorder

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients With Borderline Personality Disorder</th>
<th>Patients With Dysthymic Disorder</th>
<th>Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Global Assessment of Functioning</td>
<td>90</td>
<td>47.64</td>
<td>9.68</td>
</tr>
<tr>
<td>Personality functioning</td>
<td>90</td>
<td>3.27</td>
<td>0.82</td>
</tr>
<tr>
<td>Quality of relationships</td>
<td>90</td>
<td>2.73</td>
<td>1.06</td>
</tr>
<tr>
<td>Number of confidants</td>
<td>90</td>
<td>1.90</td>
<td>3.24</td>
</tr>
<tr>
<td>Employment stability</td>
<td>89</td>
<td>3.07</td>
<td>1.32</td>
</tr>
<tr>
<td>Current physical health</td>
<td>90</td>
<td>4.87</td>
<td>1.52</td>
</tr>
<tr>
<td>How often ill</td>
<td>90</td>
<td>3.16</td>
<td>1.49</td>
</tr>
</tbody>
</table>

* Equal variances were assumed except where Levine test showed them to be unequal at p<0.05. All degrees of freedom with decimal values are for comparisons for which unequal variance was assumed.
*p<0.01. **p<0.001.

community or whether samples of borderline personality disorder patients from hospitals and university clinics provide an overly pathological portrait of the disorder. The results point to two primary conclusions.

First, the data from prior studies, which could be expected to oversample the more severe end of the borderline spectrum, nevertheless generalized well to the patients seen by randomly selected clinicians across a wide variety of settings. In terms of axis I comorbidity, the borderline personality disorder sample in our study was very similar to other borderline personality disorder samples reviewed by Gunderson (20), with a profile of high emotional distress (in the form of mood and anxiety disorders, including posttraumatic stress disorder) and problematic ways of managing it (e.g., dissociative disorders, substance abuse, eating disorders). The findings were also consistent with Gunderson's review of research on axis II comorbidity in borderline personality disorder (20, 21, 47), in which 90%–97% of borderline personality disorder patients were found to meet the criteria for other DSM personality disorder diagnoses. In our sample, the axis II checklist identified substantial rates of comorbidity for every axis II personality disorder, with particularly high frequencies of avoidant, paranoid, and dependent personality disorders. The slightly higher rates of avoidant and dependent personality disorder diagnoses in this sample may reflect the possibility that the community treatment-seeking sample is more withdrawn and dysphoric than the more acute hospital or university clinic samples or may reflect changes in borderline personality disorder criteria between DSM-III-R and DSM-IV.

With respect to adaptive functioning, like the borderline personality disorder patients studied in prior investigations, the borderline personality disorder patients treated in everyday practice showed substantial deficits. More than two-thirds had attempted suicide, more than one-half had self-injured (mostly by cutting), and almost two-thirds had attempted suicide, more than one-half had self-injured (mostly by cutting), and almost two-thirds had been hospitalized at least once. Clinicians re-

---

**TABLE 6. Rank Order of Shedler-Westen Assessment Procedure-200 Items Not Captured by DSM-IV Borderline Personality Disorder Criteria in Current (N=90) and Previous (N=42) Samples of Borderline Personality Disorder Patients**

<table>
<thead>
<tr>
<th>Item</th>
<th>Current Sample Rank</th>
<th>Previous Sample Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tends to feel unhappy, depressed, or despondent.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Emotions tend to spiral out of control, leading to extremes of anxiety, sadness, rage, excitement, etc.</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Is unable to soothe or comfort self when distressed; requires involvement of another person to help regulate affect.</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Tends to be anxious.</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Tends to feel she/he is inadequate, inferior, or a failure.</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Tends to become irrational when strong emotions are stirred up; may show a noticeable decline from customary level of functioning.</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>

---

*a* Items are rank-ordered (from 1 to 200) according to their mean item rankings (from 0 to 7) in the current sample (see Table 5) and are listed in descending order of diagnostic importance.

*b* From a 1999 study by Westen and Shedler (23).
ported that borderline personality disorder patients had lower quality of relationships and unstable work histories, with more than one-half having lost a job in the past 5 years because of interpersonal problems.

Of particular interest are findings on physical health and abusive experiences in adulthood, as these areas of functioning have not received as much empirical attention as other aspects of adaptation. Borderline personality disorder patients appeared to have frequent or chronic minor illnesses, leading to missed appointments, days off from work, visits to the doctor, or subjective distress. Forty percent of the borderline personality disorder patients in this sample also had been in abusive relationships in adulthood, with virtually all being victimized (whether or not they also at times perpetrated violence). Nearly one-third of the borderline personality disorder patients had been the victim of rape or other serious sexual assault in adulthood, with such experiences typically occurring multiple times. A recent study by Zanarini et al. (46) produced similar results.

Second, the data provide a portrait of the average patient with DSM-IV-diagnosed borderline personality disorder seen in clinical practice, and this portrait converges with the DSM-IV description in multiple respects but diverges in others. Of the 200 items in the SWAP-200, several designed to reflect DSM-IV diagnostic criteria appeared empirically among the 20 items that were most descriptive of patients who receive a borderline personality disorder diagnosis. These items describe rejection/abandonment fears, unstable relationships, unstable identity, impulsivity, labile emotions, feelings of emptiness or boredom, and intense anger. On the other hand, several DSM-IV criteria were not ranked highly among the SWAP-200 descriptors of actual borderline personality disorder patients, and other items not included in DSM-IV received higher rankings on average than many current criteria.

Among the DSM-IV criteria that did not receive high average rankings were items describing the tendency to see people as “all good” or “all bad” (either because empirically these concepts are not as central to the diagnosis or because splitting them into two items may have decreased the ranking of both) and items describing specific forms of impulsivity, such as alcohol abuse and promiscuous sex. The hybrid criterion added to DSM-IV regarding transient psychotic symptoms and dissociative episodes did not rank highly. However, a related item that ranked tenth in the borderline personality disorder sample better appears to capture an aspect of the construct originally intended by Gunderson (21) and should be considered a replacement criterion for DSM-V: “tends to become irrational when strong emotions are stirred; may show a noticeable decline from customary level of functioning.” Although clinicians reported significant self-injury and suicidality (on the Clinical Data Form), SWAP-200 items for these phenomena did not rank as highly as other items. These differences may reflect a sampling difference between prior studies and the present investigation or may reflect the differential salience of particular high-risk behaviors when reported to an interviewer on a single occasion, typically when the patient is most symptomatic and presents at a hospital or clinic for treatment, versus when the behavior is contextualized within a longitudinal portrait of the patient over time.

Several other SWAP-200 items emerged as more descriptive than some of the current borderline personality disorder criteria, despite the fact that we used DSM-IV criteria to define the patient sample. Perhaps most important was a set of items reflecting chronic (rather than transient) aspects of emotional experience, namely the tendency “to feel unhappy, depressed, despondent” (the item with the highest ranking of all the items in the SWAP-200 composite borderline personality disorder profile) and to feel anxious. These data suggest that DSM-IV may underestimate the pain and dysphoria borderline personality disorder patients feel. They also support the view that negative affect is a central trait in borderline personality disorder as currently defined (48, 49).

Another set of items that may not receive adequate representation among the criteria for borderline personality disorder in DSM-IV describes the related trait of emotion dysregulation (on the difference between negative affect and emotion dysregulation, see references 28, 34, 40). These items describe a person whose emotions tend to spiral out of control; who can become irrational when strong emotions are stirred; who tends to “catastrophize,” seeing problems as disastrous and insolvable; and who has difficulty self-soothing and hence may become overly dependent on others to help regulate emotion. These descriptions of emotion dysregulation appear clinically richer and more specific than the DSM-IV description of “affective instability due to a marked reactivity of mood.” All of these items distinguished borderline personality disorder patients from dysthymic disorder patients, who share the trait of negative affect, and occurred at a similarly high rate in previous efforts to develop an empirical prototype of borderline personality disorder (22).

Limitations and Potential Objections

The data from this study, like the data from a number of studies from our laboratory (e.g., references 22, 25, 27, 29), point to the potential utility of practice research network methods in research on personality disorders. The convergence of multiple informants would clearly be ideal (although most studies of psychopathology rely on a single observer—the patient—by means of either self-report or structured interviews); however, data from experienced clinical observers who interact with the patient over time and hence can provide a longitudinal portrait provide a complementary standpoint to that typically seen in psychiatric research.

Another set of limitations concerns the makeup of the current sample. Psychologists were disproportionately...
represented in the sample, relative to psychiatrists (80% and 20%, respectively), and the overall response rate was relatively low, compared to our previous studies. Although we cannot be sure that some unknown bias was not introduced by clinicians’ decisions to participate or not participate, the data provided by psychologists and by psychiatrists did not show any pattern of differences in this or any of our prior studies using this method, despite substantially different response rates, and our findings converged with those of research from medical centers that used completely different sampling methods. The similarities in these findings suggest that such biases are not likely substantial. Limiting the study to patients in psychotherapy also introduced the possibility that we were oversampling higher-functioning borderline personality disorder patients. However, the data suggested otherwise: The majority of the borderline personality disorder patients in the study reported histories of psychiatric hospitalizations, suicide attempts, and self-injurious behavior, and their average GAF score (mean=47.64, SD=9.68) indicated serious impairment. Finally, male and non-Caucasian borderline personality disorder patients are understudied groups, and broader sampling, including oversampling to maximize representativeness of the population, would strengthen future investigations.

References

The Changing Prevalence and Severity of Obsessive-Compulsive Disorder Criteria From DSM-III to DSM-IV

Rocco Crino, Ph.D.
Tim Slade, Ph.D.
Gavin Andrews, M.D.

Objective: Relative to other mental disorders, the prevalence of obsessive-compulsive disorder (OCD) in the general population is not well established. Some epidemiological surveys have determined the prevalence of DSM-III OCD, but this is one of the first reports, to the authors’ knowledge, of DSM-IV OCD’s prevalence.

Method: Data from the Australian National Survey of Mental Health and Well-Being, a nationally representative epidemiological survey of mental disorders, were analyzed. The prevalence and associated characteristics of DSM-IV OCD were identified, and then the data were rescored for DSM-III OCD. Cases defined by each system were compared.

Results: The 12-month prevalence of DSM-IV OCD was 0.6%, considerably less than found in surveys employing DSM-III diagnostic criteria. DSM-IV OCD showed significantly higher levels of comorbidity, disability, health service use, and treatment received.

Conclusions: Changes in the reported prevalence and severity of OCD between DSM-III and DSM-IV cases are most likely a function of the differences in diagnostic criteria between DSM-III and DSM-IV.

Obsessive-compulsive disorder (OCD) is a chronic and debilitating condition of which the prevalence in the general population remains somewhat controversial. Early investigations suggested that OCD was a relatively rare condition, with Black (1) reporting that only 3% of all neurotic patients from a series of studies were diagnosed as obsessional and Woodruff and Pitts (2) estimating that OCD affects a mere 0.05% of the population. Karno et al. (3) reported the lifetime prevalence rate of DSM-III OCD in the Epidemiologic Catchment Area (ECA) studies as 2.5%, averaged across five U.S. catchment areas. The point prevalence (2 weeks) was 1.2%, and the annual prevalence was 1.6%. The Cross National Collaborative Group (4) summarized the findings from the various DSM-III studies and reported lifetime prevalence rates in seven countries as ranging from a low of 0.7% (in Taiwan) to a high of 2.5% (in Puerto Rico). Taiwan excluded, the 12-month prevalence rates across the other sites were consistent; the U.S. study reported 1.3%; Edmonton, Canada 1.4%; Puerto Rico 1.8%; Munich 1.6%; and Korea and New Zealand 1.1%.

The findings of higher-than-expected rates of OCD in epidemiological studies resulted in OCD being labeled a “hidden epidemic” (5) and being ranked 20th in the Global Burden of Disease studies among all diseases as a cause of disability-adjusted life years lost in developed countries (6). Subsequent investigations cast some doubt on the reliability of the results obtained in these epidemiological studies. Nelson and Rice (7) examined the temporal stability of the lifetime National Institute of Mental Health Diagnostic Interview Schedule (DIS) diagnosis of DSM-III OCD by reinterviewing ECA subjects (wave 1) with a positive diagnosis of OCD 12 months later (wave 2). The authors reported a kappa statistic of 0.2, with only 19% of the original 291 subjects reporting lifetime OCD at wave 1 receiving a similar diagnosis at wave 2. Because of these findings, the authors concluded that the wave 1 lifetime prevalence may have been largely made up of false positives and that the validity of DIS-diagnosed OCD is questionable. Stein et al. (8) also cast doubt on the validity of the DSM-III epidemiological data. The authors examined the prevalence of DSM-IV OCD in 2,261 Canadian subjects who were interviewed by lay interviewers over the telephone. The study used a modified version of the Composite International Diagnostic Interview and noted a 1-month prevalence rate of 3.1%. However, when a subgroup of positively diagnosed OCD subjects was reinterviewed by clinicians using a structured clinical interview (the Structured Clinical Interview for DSM-IV), the 1-month prevalence rate fell to 0.6%, with a further 0.6% fulfilling the authors’ criteria for subclinical OCD. The main source of discrepancy was the tendency for respondents to label everyday sources of “worry, concern, preoccupation or interests as obsessions” (8, p. 1123), which lay interviewers were unable to differentiate from true obsessions, as well as a tendency for respondents to overestimate the degree of disability or distress associated with the symptoms.

Up to 80% of the general population may experience intrusive, unpleasant, unwanted thoughts similar to those seen in OCD (9, 10), albeit to a less distressing and frequent degree than in OCD. More than half the population may engage in ritualized behavior (11), also to a less distressing degree than clinical subjects. The possibility that OCD diag-
nosed with the DIS/Composite International Diagnostic Interview is overinclusive because of its inability to determine the frequency, distress, and disability of obsessive-compulsive symptoms is not surprising. Similarly, the finding that everyday worries were often confounded with obsessions is not surprising, particularly because one item that makes up the diagnosis of OCD in the DIS and the Composite International Diagnostic Interview (version 1.1) includes “persistent or unpleasant thoughts that relatives who are away have been hurt or killed,” a symptom more in keeping with generalized anxiety disorder than with OCD.

Possible inaccuracies in the reported DSM-III prevalence rates were also noted by Korn and Golding (12), who suggested that more refined definitions of obsessions and compulsions associated with DSM-III-R would “reduce the number of ‘borderline’ cases in future epidemiological studies employing the revised criteria and will produce a slight lowering of prevalence rates” (12, p. 207). Changes in the DSM-III-R and DSM-IV diagnostic criteria have indeed focused on better defining obsessions and compulsions while also emphasizing the degree of distress and impairment required for the diagnosis to be positive. For example, DSM-III suggests that “obsessions or compulsions are a significant source of distress to the individual or interfere with social or role functioning” (p. 235). DSM-IV notes that the obsessions or compulsions should be recognized at some point as “excessive or unreasonable” and that they “cause marked distress, are time-consuming, or significantly interfere with the person’s normal routine, occupational (or academic) functioning or usual social activities or relationships” (p. 423). Of equal importance, DSM-IV specifically excludes other axis I disorders that may confound the diagnosis. Thus, the prevalence of OCD may also be affected by the changes in the diagnostic criteria from DSM-III to DSM-IV.

The purpose of the current study was twofold: first, to report on the prevalence of DSM-IV OCD in Australia, as determined in the Australian National Survey of Mental Health and Well-Being, and second, to examine the prevalence rates when the data are rescored according to the DSM-III diagnostic criteria. This was the first such analysis in a nationally representative survey.

Method

Survey Design and Sample

The Australian National Survey of Mental Health and Well-Being was a nationwide household survey of adults conducted in 1997. It aimed to determine the prevalence of both ICD-10 and DSM-IV mental disorders in the community and to describe their associated disability and service use. The overall method and design of the survey have been described in detail elsewhere (13). The interview was computerized and conducted by trained interviewers from the Australian Bureau of Statistics, a statutory body responsible for conducting such surveys using ethical protocols that include written informed consent. The Australian Bureau of Statistics employed a stratified multistage sampling process resulting in a sample of 10,641 persons over the age of 18, a response rate of 78.1%. The sample was weighted to conform to the age and sex distribution of the Australian population and to account for the probability of selection.

Assessment of Diagnosis

DSM-IV diagnoses of OCD and other anxiety, affective, and substance use disorders were made by using the Composite International Diagnostic Interview (version 2.1). The presence of DSM-IV obsessions was assessed with 10 separate questions. The first two asked about unpleasant, intrusive, inappropriate, and persistent thoughts, such as “the idea that your hands are dirty or have germs on them” or “the idea that you might harm someone, even though you really didn’t want to.” If either of these questions was answered positively, eight remaining questions asked about attempts to suppress the thoughts; the excessiveness, unreasonableness, distress, and interference in normal functioning caused by the thoughts; whether the thoughts were time-consuming; and whether they were exclusively related to the symptoms of a comorbid axis I disorder. All symptoms were assessed for their presence in the 12 months before the interview. One-month diagnoses were derived from 12-month diagnoses by restricting the sample to those who experienced the symptoms in the 4 weeks before the interview. Personality disorder diagnoses were made with the use of a screening questionnaire for personality disorders (14). This screening questionnaire was not a fully standardized diagnostic interview for personality disorders but was included in the interview to provide a tentative estimate of the prevalence of personality disturbance.

Assessment of Disability

Three measures of disability are reported. The Medical Outcomes Study 12-Item Short-Form Health Survey (15) is a measure of disablement. It has two regression-weighted scales: a mental health summary scale and a physical health summary scale. The continuous scales are scored such that the mean is 50 and the standard deviation is 10. Higher scores indicate less disability. The role functioning scale of the Brief Disability Questionnaire (16) assesses the extent to which physical and social role activities have been limited in the 4 weeks before the interview. The Disability Days Scale (17) is a summary measure of the number of days in the past 4 weeks that an individual has been unable to perform or has had to cut down on his or her normal activities because of ill health. For ease of interpretation, scores on the three continuous measures of disability were transformed into z scores (calculated as the individual score minus the mean of all scores divided by the standard deviation of all scores). Thus, when these measures are used as independent variables in logistic regression analyses, the resultant odds ratios represent a one standard deviation—as opposed to a 1-point—shift in the disability measure.

Assessment of Service Use and Treatment Received

The respondents were asked about both inpatient and outpatient service use in the 12 months before the interview. If the respondents had consulted any health professional, they were asked how many times they had visited the office for mental health problems, including “stress, anxiety, depression, or de-
results

Prevalence and Demographic Correlates of DSM-IV OCD

The weighted prevalence of 12- and 1-month DSM-IV OCD by age and sex is shown in Table 1. A total of 81 subjects (0.6%) were assigned a 12-month diagnosis of OCD. The weighted 1-month prevalence was 0.5%. The average age at onset was 27 years (SE=4.5) for men and 25 years (SE=3.1) for women.

Unadjusted and adjusted demographic correlates of 12-month DSM-IV OCD were calculated. No differences between people with OCD and the remainder of the sample were noted for sex, marital status, education, language spoken at home, migration status, urbanicity, or household composition. There was a significantly higher prevalence of OCD among those who were unemployed and those who were not in the labor force than among those who were employed (adjusted odds ratio=8.7 and 4.3, respectively), and older individuals (>55 years) were significantly less likely than those in younger age groups to have OCD (adjusted odds ratio=0.1).

Comorbidity

Comorbidity was common. The percents of DSM-IV OCD subjects who also met criteria for other DSM-IV disorders are shown in Table 2. Overall, individuals with OCD were significantly more likely than people without OCD to have met criteria for at least one affective, anxiety, substance use, or personality disorder (odds ratio=16.4), with 79.7% of those with OCD having another disorder. OCD subjects were also significantly more likely to have two (odds ratio=10.6) and three or more other disorders (odds ratio=57.8) than those without a diagnosis of OCD, with 46% of individuals with OCD meeting the criteria for three or more disorders in the 12 months before the interview.

Among the individual disorders, comorbidity with major depression was highest at 54% (odds ratio=5.4). Although a general association between OCD and any other anxiety disorder was found (odds ratio=5.4), the presence of OCD was significantly associated only with the presence of panic disorder (odds ratio=3.9) and posttraumatic stress disorder (odds ratio=2.4). No relationship was noted between OCD and substance use disorders. Although there was a more general association between OCD and any personality disorder (odds ratio=3.3), there was no specific association between OCD and anankastic personality disorder, the ICD-10 personality disorder that is most similar to DSM-IV obsessive-compulsive personality disorder.

Disability, Health Service Use, and Treatment Received

Disability, health service use, and treatment received for DSM-IV OCD are shown in Table 3. Compared to the rest of the population, OCD was associated with significant disability, as measured by the mental component scale of the 12-item Short-Form Health Survey (odds ratio=0.5; low scores indicate high levels of disability), the role functioning scale of the Brief Disability Questionnaire (odds ratio=1.6), and the Disability Days Scale (odds ratio=1.3).

The proportion of OCD subjects who consulted health professionals in the 12 months before the interview was consistent with the levels of disability. Fifty-nine percent of those with OCD consulted at least one health professional for mental health problems were asked about the type of help received.

Analysis

Odds ratios were derived from logistic regression models and are presented as either unadjusted odds ratios derived from bivariate regression models where only one predictor was entered into the model or adjusted odds ratios where some or all predictors were entered into the model at the same time to obtain estimates that account for the effects of other variables. Standard errors around proportions and confidence intervals around odds ratios were calculated with delete-one jackknife repeated replications (18) to account for the complex sampling design. The SUDAAN software package, designed specifically for use with complex survey samples, was used for all analyses (19).

Rescoring to Obtain DSM-III Diagnoses of OCD

Data were rescored to generate DSM-III OCD diagnoses. This was made possible by the similarity between the items in the Composite International Diagnostic Interview and the DIS that assess OCD symptoms as well as the method that both instruments employed to probe for clinical significance. To meet criteria for DSM-III OCD, the respondents were required to meet two criteria. First, a positive response on either of the two questions about unpleasant, intrusive, inappropriate, and persistent thoughts was needed, and then a positive response on one of the three questions about engaging in repetitive behaviors was required (DSM-III criterion A). The question about cognitive compulsions was not included because it was not considered part of the DSM-III diagnosis of OCD. Second, a positive response to the questions assessing distress and/or interference was required (DSM-III criterion B). The exclusion rules in DSM-III criterion C were not operationalized. The modifications to the scoring algorithm used in the current study are available from Dr. Slade.

Results

TABLE 1. Weighted Prevalence of DSM-IV Obsessive-Compulsive Disorder by Age and Sex in the Australian National Survey of Mental Health and Well-Being

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Men (N=64)</th>
<th>Women (N=64)</th>
<th>Total (N=128)</th>
<th>Men (N=81)</th>
<th>Women (N=81)</th>
<th>Total (N=162)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>SE</td>
<td>%</td>
<td>SE</td>
<td>%</td>
<td>SE</td>
</tr>
<tr>
<td>18–34</td>
<td>0.6</td>
<td>0.1</td>
<td>0.7</td>
<td>0.2</td>
<td>0.6</td>
<td>0.1</td>
</tr>
<tr>
<td>35–54</td>
<td>1.0</td>
<td>0.3</td>
<td>0.9</td>
<td>0.2</td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td>≥55</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>Total</td>
<td>0.6</td>
<td>0.1</td>
<td>0.7</td>
<td>0.1</td>
<td>0.6</td>
<td>0.1</td>
</tr>
</tbody>
</table>
for a mental health problem, although many consulted more than one professional. A general practitioner was consulted by 53.2% of those with OCD and a psychologist or psychiatrist by 28.6% of those with OCD. A summary of the treatments received indicates that approximately half of the OCD subjects who consulted any health professional for a mental health problem received treatment of any kind, and 42% received an evidence-based intervention.

**Comparison of DSM-III and DSM-IV OCD**

The overlap between the prevalence of DSM-IV and DSM-III OCD is shown in Table 4. The weighted prevalence of 12-month DSM-III OCD was 2.1%, a prevalence rate somewhat higher than the rates reported in the cross-national sites but not markedly higher than Puerto Rico at 1.8% and Munich at 1.6%. It is possible for individuals to meet both DSM-IV and DSM-III diagnostic criteria for OCD. For statistical comparison purposes, two mutually exclusive groups were constructed, one containing all DSM-IV cases of OCD (N=81) and the other containing DSM-III but not DSM-IV cases of OCD (N=172). These two groups represent all DSM-III or DSM-IV OCD cases in the sample. It should be noted that 89% of those with DSM-IV OCD also met the criteria for DSM-III OCD.

Comparisons between the DSM-IV and DSM-III cases of OCD are displayed in Table 5. DSM-IV subjects of OCD were more severe on a number of levels. They were significantly more likely among the 35–54-year-old age group compared to the younger age group as well as among the unemployed compared to the employed. DSM-IV OCD subjects were more likely to have comorbid psychiatric disorders except for substance use disorders (statistically significant odds ratios ranged from 2.4 to 4.7). With control for the presence of any comorbid disorder, DSM-IV OCD subjects were also more likely than DSM-III cases to be disabled on two out of the three disability measures, and they were more likely to use health services and to receive treatment.

**Discussion**

The 12-month prevalence of DSM-IV OCD in Australia was found to be 0.6% of the adult population, less than that reported in epidemiological studies employing DSM-III criteria. There are a number of general similarities between the current study and previous epidemiological investigations. First, the average age at onset of OCD was 26.1 years, similar to the age at onset reported in the U.S. ECA survey (26.2 years) and in Edmonton and New Zealand (27.2 years) (4). Second, the sex ratio of close to one is similar to the majority of other epidemiological studies. Third, there was a noteworthy decrease in prevalence of OCD in older age groups (>55 years). Fourth, the high rate of OCD comorbidity with anxiety disorders and affective disorders is noted in a number of epidemiological and clinical studies. Overall, lower rates of alcohol and substance abuse were found in the current 12-month Aus-

<table>
<thead>
<tr>
<th>Comorbid Disorders</th>
<th>%</th>
<th>SE</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: individual disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major depression</td>
<td>53.7</td>
<td>5.8</td>
<td>5.4**</td>
<td>2.9–10.1</td>
</tr>
<tr>
<td>Dysthymia</td>
<td>13.8</td>
<td>5.0</td>
<td>1.2</td>
<td>0.4–3.8</td>
</tr>
<tr>
<td>Panic with or without agoraphobia</td>
<td>28.6</td>
<td>5.0</td>
<td>3.9*</td>
<td>1.9–8.1</td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>9.2</td>
<td>2.9</td>
<td>1.5</td>
<td>0.5–3.9</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>36.8</td>
<td>5.9</td>
<td>2.0</td>
<td>0.8–4.9</td>
</tr>
<tr>
<td>Social phobia</td>
<td>27.0</td>
<td>5.4</td>
<td>1.6</td>
<td>0.6–4.6</td>
</tr>
<tr>
<td>Posttraumatic stress disorder</td>
<td>23.1</td>
<td>4.1</td>
<td>2.4*</td>
<td>1.1–5.4</td>
</tr>
<tr>
<td>Alcohol abuse or dependence</td>
<td>15.1</td>
<td>4.4</td>
<td>0.9</td>
<td>0.4–2.1</td>
</tr>
<tr>
<td>Drug abuse or dependence</td>
<td>12.7</td>
<td>3.6</td>
<td>1.5</td>
<td>0.6–3.7</td>
</tr>
<tr>
<td>Anankastic personality disorder</td>
<td>22.8</td>
<td>4.4</td>
<td>1.6</td>
<td>0.7–3.4</td>
</tr>
<tr>
<td>Any other personality disorder</td>
<td>41.6</td>
<td>5.5</td>
<td>2.4</td>
<td>1.0–5.9</td>
</tr>
<tr>
<td>Model 2: disorder groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any affective disorder</td>
<td>55.4</td>
<td>6.1</td>
<td>4.7**</td>
<td>2.5–8.9</td>
</tr>
<tr>
<td>Any anxiety disorder</td>
<td>60.7</td>
<td>5.1</td>
<td>5.4**</td>
<td>2.7–10.7</td>
</tr>
<tr>
<td>Any substance use disorder</td>
<td>23.9</td>
<td>4.0</td>
<td>1.2</td>
<td>0.7–2.2</td>
</tr>
<tr>
<td>Any personality disorder</td>
<td>47.7</td>
<td>5.6</td>
<td>3.3*</td>
<td>1.7–6.3</td>
</tr>
<tr>
<td>Model 3: number of other disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No other disorders</td>
<td>20.3</td>
<td>5.8</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>One other disorder</td>
<td>23.7</td>
<td>5.3</td>
<td>7.6**</td>
<td>2.9–20.4</td>
</tr>
<tr>
<td>Two other disorders</td>
<td>10.1</td>
<td>3.9</td>
<td>10.6*</td>
<td>3.1–36.4</td>
</tr>
<tr>
<td>Three or more other disorders</td>
<td>46.0</td>
<td>5.3</td>
<td>57.8**</td>
<td>29.1–114.5</td>
</tr>
<tr>
<td>Model 4: any other disorder</td>
<td>79.7</td>
<td>5.8</td>
<td>16.4*</td>
<td>7.5–36.1</td>
</tr>
</tbody>
</table>

* Odds ratios were calculated by using parameter estimates from logistic regression models. They represented the odds of having each disorder (single or group) for persons with OCD compared to those without OCD. The odds ratios were derived from four different models: model 1=each single mental disorder controlled for the presence of any other single mental disorder; model 2=each disorder group controlled for the presence of any other disorder group; model 3=a summary measure of the number of mental disorders where no mental disorders was the reference category; and model 4=any disorder.

**OCD plus any personality disorder other than anankastic.**

**OCD plus any anxiety disorder other than OCD.**

**OCD plus any disorder other than OCD.**

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

The Australian cohort had 17.6% of those with 12-month DSM-IV OCD in the current investigation, compared to a DSM-III lifetime rate of 17.6% of the U.S. ECA survey OCD cohort and 26.5% in Edmonton (12, 20).

There are a number of differences. In general, comorbidity rates of OCD with affective and anxiety disorders are higher in the current study than in previous investigations. For example, in the U.S. ECA survey, Kanno et al. (3) reported that 31.7% of the OCD cohort met criteria for lifetime major depression, while 13.8% met the criteria for panic disorder. Similar rates of lifetime depression (29.6%) and panic (9.8%) were reported in Edmonton (20). In contrast, 12-month major depression was reported in 53.7% of the OCD subjects in the current study, and panic (with and without agoraphobia) was found in 28.6% of the OCD cohort. The fact that the current investigation focuses on 12-month diagnoses while the U.S. ECA survey focused on lifetime diagnoses would suggest the opposite, i.e., that there would be a higher rate of comorbidity in the lifetime studi-
TABLE 3. Disability, Health Service Use, and Treatment Received Among Persons With 12-Month DSM-IV Obsessive-Compulsive Disorder

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SE</th>
<th>%</th>
<th>SE</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental component scale of the Medical Outcomes Study</td>
<td>37.3</td>
<td>2.1</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4–0.7</td>
<td></td>
</tr>
<tr>
<td>Role functioning scale of the Brief Disability Questionnaire</td>
<td>3.7</td>
<td>0.4</td>
<td>1.6</td>
<td>1.4</td>
<td>1.4–1.9</td>
<td></td>
</tr>
<tr>
<td>Disability Days Scale</td>
<td>8.3</td>
<td>1.4</td>
<td>1.3</td>
<td>1.1</td>
<td>1.1–1.6</td>
<td></td>
</tr>
<tr>
<td>Health service use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any health professional</td>
<td>59.3</td>
<td>5.1</td>
<td>4.5</td>
<td>3.1</td>
<td>3.1–6.5</td>
<td></td>
</tr>
<tr>
<td>General practitioner</td>
<td>53.2</td>
<td>6.2</td>
<td>4.9</td>
<td>3.2</td>
<td>3.2–7.5</td>
<td></td>
</tr>
<tr>
<td>Mental health professional</td>
<td>28.6</td>
<td>6.5</td>
<td>4.5</td>
<td>2.4</td>
<td>2.4–8.6</td>
<td></td>
</tr>
<tr>
<td>Treatment received</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any treatment</td>
<td>55.1</td>
<td>5.8</td>
<td>4.3</td>
<td>2.7</td>
<td>2.7–6.7</td>
<td></td>
</tr>
<tr>
<td>Information about illness</td>
<td>30.2</td>
<td>6.2</td>
<td>4.9</td>
<td>2.8</td>
<td>2.8–9.9</td>
<td></td>
</tr>
<tr>
<td>Evidence-based treatment</td>
<td>42.3</td>
<td>5.4</td>
<td>3.5</td>
<td>2.3</td>
<td>2.3–5.6</td>
<td></td>
</tr>
<tr>
<td>Nonspecific counseling</td>
<td>41.7</td>
<td>5.5</td>
<td>4.0</td>
<td>2.6</td>
<td>2.6–6.2</td>
<td></td>
</tr>
<tr>
<td>Social support</td>
<td>18.3</td>
<td>4.6</td>
<td>4.1</td>
<td>2.2</td>
<td>2.2–8.0</td>
<td></td>
</tr>
</tbody>
</table>

*a Odds ratios were calculated with parameter estimates from separate logistic regression models with control for the presence of any mental disorder. Odds ratios for the three continuous measures of disability represented a one standard deviation shift in scores on the respective disability measure.

b Psychiatrist, psychologist, or mental health team.

c Medicines, tablets, or cognitive behavior therapy.

d Psychotherapy or counseling.

e Help with house and money problems, help with the person’s ability to work, help with looking after themselves, help with meeting people, or any other kind of help.

*p<0.05. **p<0.001.

TABLE 4. Overlap Between Prevalence of 12-Month DSM-IV and DSM-III Obsessive-Compulsive Disorder (OCD)

<table>
<thead>
<tr>
<th>DSM-III OCD</th>
<th>DSM-IV OCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>Negative</td>
<td>10,388</td>
</tr>
<tr>
<td>Positive</td>
<td>172</td>
</tr>
<tr>
<td>Total</td>
<td>81 (0.6%)</td>
</tr>
</tbody>
</table>
of the disorder, and this in itself may be problematic in de-
termining the community prevalence of OCD.

Clearly, clinical presentation has informed changes in
the diagnostic criteria. While benefits have included bet-
ter definition and demarcation of obsessions and compul-
sions, it is more than likely that the more severely disabled
are seen for treatment, resulting in an artificially raised se-
verity threshold for a positive diagnosis. As noted in previ-
ous sections, obsessions and compulsions are common in
the community, although reported disability and interfer-
ence are less common. Ratings of disability and interfer-
ence are often subjective and do not take into account
such influencing factors as accommodation of symptoms,
embarrassment, or denial of interference. Although the
evidence is merely anecdotal, it is not uncommon for pa-
tients to be seen for treatment because of family or a sig-
nificant other’s insistence and to deny any interference or
disability. Similarly, it is not uncommon for loved ones and
health professionals to seek assistance for OCD sufferers
who refuse to acknowledge any difficulty as they accom-
mmodate their symptoms (e.g., getting up an hour early to
perform checking rituals before leaving home). Such com-
community cases, the prevalence of which is unknown, would
not have been included in the current survey.

Other issues in the identification of positive cases also
need to be considered in interpreting the data from the cur-
rent study. Any structured instrument administered by lay
interviewers will be less reliable in properly assessing the
sometimes complex presentation of OCD, and clinical as-
sessment is sometimes necessary to differentiate it from
other disorders where there may be overlap. Clinical reas-
sessment of a proportion of identified DSM-III and DSM-IV
OCD cases would have considerably strengthened the find-
ings of the current study; however, privacy agreements pre-
cluded repeated access to any participants in the survey. Al-
ternatively, the addition of self-report obsessive-compulsive

### TABLE 5. Demographic and Clinical Characteristics of Persons With 12-Month DSM-III (But Not DSM-IV) Obsessive-Compulsive Disorder (OCD) and a Comparison Between Persons With DSM-IV and DSM-III OCD

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>OCD (DSM-III But Not DSM-IV)</th>
<th>DSM-IV Versus DSM-III OCD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>SE</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–34</td>
<td>53.5</td>
<td>4.6</td>
</tr>
<tr>
<td>35–54</td>
<td>29.8</td>
<td>4.9</td>
</tr>
<tr>
<td>≥55</td>
<td>16.7</td>
<td>2.6</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time or part-time</td>
<td>54.4</td>
<td>4.5</td>
</tr>
<tr>
<td>Unemployed</td>
<td>5.5</td>
<td>1.7</td>
</tr>
<tr>
<td>Not in the labor force</td>
<td>40.1</td>
<td>4.1</td>
</tr>
<tr>
<td>Comorbid disorders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any affective disorder</td>
<td>20.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Any anxiety disorder&lt;sup&gt;b&lt;/sup&gt;</td>
<td>36.2</td>
<td>4.9</td>
</tr>
<tr>
<td>Any substance use disorder</td>
<td>17.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Any personality disorder&lt;sup&gt;c&lt;/sup&gt;</td>
<td>27.9</td>
<td>3.8</td>
</tr>
<tr>
<td>Any mental disorder</td>
<td>57.4</td>
<td>5.4</td>
</tr>
<tr>
<td>Disability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental component scale of the Medical Outcomes Study</td>
<td>45.8</td>
<td>1.0</td>
</tr>
<tr>
<td>12-Item Short-Form Health Survey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role functioning scale of the Brief Disability Questi</td>
<td>2.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Disability Days Scale</td>
<td>5.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Health service use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any health professional</td>
<td>24.7</td>
<td>3.9</td>
</tr>
<tr>
<td>General practitioner</td>
<td>19.9</td>
<td>3.5</td>
</tr>
<tr>
<td>Any mental health professional&lt;sup&gt;d&lt;/sup&gt;</td>
<td>6.7</td>
<td>2.1</td>
</tr>
<tr>
<td>Treatment received</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any treatment</td>
<td>23.0</td>
<td>3.9</td>
</tr>
<tr>
<td>Information about illness</td>
<td>10.2</td>
<td>2.5</td>
</tr>
<tr>
<td>Evidence-based treatment&lt;sup&gt;e&lt;/sup&gt;</td>
<td>19.6</td>
<td>3.5</td>
</tr>
<tr>
<td>Nonspecific counseling</td>
<td>13.6</td>
<td>2.5</td>
</tr>
<tr>
<td>Social support&lt;sup&gt;f&lt;/sup&gt;</td>
<td>5.9</td>
<td>1.5</td>
</tr>
</tbody>
</table>

<sup>a</sup> Odds ratios were calculated by using parameter estimates from separate logistic regression models. Odds ratios for the three continuous measures of disability represent a one standard deviation shift in scores on the respective disability measure.

<sup>b</sup> OCD plus any anxiety disorder other than OCD.

<sup>c</sup> OCD plus any personality disorder other than anankastic.

<sup>d</sup> Psychiatrist, psychologist, or mental health team.

<sup>e</sup> Medicines or cognitive behavior therapy.

<sup>f</sup> Psychotherapy or counseling.

<sup>g</sup> Help with house and money problems, help with the person’s ability to work, help with looking after themselves, help with meeting people, or any other kind of help.

*p<0.05   **p<0.001.
inventories (e.g., the Padua Inventory, the Yale-Brown Obsessive Compulsive Scale self-report) may have further assisted in identifying positive cases, particularly among those fulfilling DSM-III but not DSM-IV criteria. Of greater importance is the possibility that a significant proportion of those fulfilling DSM-III but not DSM-IV criteria may have fallen below the cutoff for DSM-IV disability and interference criteria on the basis of receiving effective treatment. One-fifth (19.6%) of those fulfilling DSM-III criteria reported receiving evidence-based treatment (pharmacotherapy or cognitive behavior therapy), and 23% reported receiving treatment of any kind. It could be argued that some or all of these individuals may have fulfilled DSM-IV criteria in the absence of such treatment, i.e., that they were treatment responders, particularly those reporting evidence-based treatment. Although 42.3% of the DSM-IV sample had received evidence-based treatment, their overall greater disability and comorbidity would suggest that many of these individuals may have been at the more severe end of the OCD spectrum and their symptoms may have been less likely to fall below the DSM-IV disability threshold, even though there may have been some treatment response. Finally, the literature suggests that a significant proportion of OCD subjects report a waxing and waning course, and a small proportion are classified as episodic (21, 22). It is possible that a proportion of those whose symptoms were classified as meeting DSM-III but not DSM-IV OCD criteria also did not meet disability and interference thresholds as a result of the natural severity fluctuations of OCD. Again, those fulfilling DSM-IV criteria, being more severely ill, may have had symptoms remaining above the thresholds, despite similar fluctuations.

Conclusions

The 12-month prevalence of DSM-IV OCD in Australia is 0.6%, a figure considerably less than expected from estimates of previous DSM-III epidemiological studies conducted in other countries. The findings from the current study indicate that the differing prevalence is a function of the changes in diagnostic criteria from DSM-III to DSM-IV, which are reflected in the instruments used to assess population prevalence. Comparison of identified DSM-IV cases with DSM-III cases that did not meet DSM-IV criteria indicated considerably greater severity among the DSM-IV cases in terms of comorbidity, disability, service use, and unemployment. Previous epidemiological estimates may have been confounded by identification of milder (non-clinical) cases and other disorders where symptoms may overlap with OCD. The cases identified in the current study are similar to those seen in clinical practice in terms of disability and comorbidity. Although identified cases are reflective of the OCD population presenting for clinical treatment, they may not be an accurate reflection of the full dimension of OCD-type symptoms in the community.

References

15. Ware J Jr, Kosinski M, Keller SD: A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996; 34:220–223
17. Kessler R, Frank G: The impact of psychiatric disorders on work loss days. Psychol Med 1997; 27:861–873

Received Dec. 28, 2003; revision received May 25, 2004; accepted Aug. 30, 2004. From the Clinical Research Unit for Anxiety and Depression, St. Vincent’s Hospital, the School of Psychiatry, University of New South Wales, Sydney. Address correspondence and research requests to Dr. Slade, Clinical Research Unit for Anxiety and Depression, St. Vincent’s Hospital, 299 Forbes St., Darlinghurst, N.S.W., 2010, Austr- alia; tims@unsw.edu.au (e-mail).

Two-Year Prevalence and Stability of Individual DSM-IV Criteria for Schizotypal, Borderline, Avoidant, and Obsessive-Compulsive Personality Disorders: Toward a Hybrid Model of Axis II Disorders

Thomas H. McGlashan, M.D.
Carlos M. Grilo, Ph.D.
Charles A. Sanislow, Ph.D.
Elizabeth Ralevski, Ph.D.
Leslie C. Morey, Ph.D.
John G. Gunderson, M.D.
Andrew E. Skodol, M.D.
M. Tracie Shea, Ph.D.
Mary C. Zanarini, Ed.D.
Donna Bender, Ph.D.
Robert L. Stout, Ph.D.
Shirley Yen, Ph.D.
Maria Pagano, Ph.D.

Objective: This study tracked the individual criteria of four DSM-IV personality disorders—borderline, schizotypal, avoidant, and obsessive-compulsive personality disorders—and how they change over 2 years.

Method: This clinical sample of patients with personality disorders was derived from the Collaborative Longitudinal Personality Disorders Study and included all participants with borderline, schizotypal, avoidant, or obsessive-compulsive personality disorder for whom complete 24-month blind follow-up assessments were obtained (N=474). The authors identified and rank-ordered criteria for each of the four personality disorders by their variation in prevalence and changeability (remission) over time.

Results: The most prevalent and least changeable criteria over 2 years were paranoid ideation and unusual experiences for schizotypal personality disorder, affective instability and anger for borderline personality disorder, feeling inadequate and feeling socially inept for avoidant personality disorder, and rigidity and problems delegating for obsessive-compulsive personality disorder. The least prevalent and most changeable criteria were odd behavior and constricted affect for schizotypal personality disorder, self-injury and behaviors defending against abandonment for borderline personality disorder, avoiding jobs that are interpersonal and avoiding potentially embarrassing situations for avoidant personality disorder, and miserly behaviors and strict moral behaviors for obsessive-compulsive personality disorder.

Conclusions: These patterns highlight that within personality disorders the relatively fixed criteria are more trait-like and attitudinal, whereas the relatively intermittent criteria are more behavioral and reactive. These patterns suggest that personality disorders are hybrids of traits and symptomatic behaviors and that the interaction of these elements over time helps determine diagnostic stability. These patterns may also inform criterion selection for DSM-V.

This article examines the individual criteria of four DSM-IV personality disorders: schizotypal, borderline, avoidant, and obsessive-compulsive. For each disorder we track how each criterion varies over a 2-year period, compared with every other criterion. We do this to rank-order them in terms of their prevalence and changeability or variance within their personality disorder categories, thus providing clues as to the presence and character of underlying dimensions or phenotypes as well as providing data about the centrality and importance of each criterion for future iterations of personality disorder nosology.

The individual criteria for the DSM-IV axis II personality disorders were first articulated in DSM-III. The criteria for borderline personality disorder largely emerged from a burgeoning clinical literature on the disorder (1–5) and from a discriminant function analysis of data from patients judged by clinicians to have borderline personality disorder, compared with samples of patients with other disorders, including schizophrenia and dysthymia (6). The criteria for schizotypal personality disorder were based on descriptions of first-degree relatives of probands with schizophrenia in Danish adoption studies (7). The criteria for most of the personality disorders, however, were proposed by clinicians on the DSM-III Task Force (8).

The selection process of criteria for the DSM-IV personality disorders was built on a database of comorbidity and criterion diagnostic efficiency studies generated by using DSM-III and DSM-III-R personality disorder categories of criteria. Data were available on the sensitivity, specificity, positive and negative predictive power, and phi coefficients of every personality disorder criterion for its own personality disorder category and for other personality disorder categories (summarized in Widiger [9]). Many of the DSM-III and DSM-III-R personality disorder criteria were retained with some revisions but were rank-ordered for DSM-IV on the basis of their importance as measured...
by their diagnostic efficiency credentials and expert clinician consensus (10).

The DSM-IV personality disorder criteria have often been described as heterogeneous entities. For example, Parker et al. (11) considered personality disorders to be an amalgam of two constructs, personality style and/or disorder. Rating personality styles and manifestations of disorder in a clinical sample of depressed patients, they found that the personality disorder criteria judged to most closely describe personality style often acted as “proxy criteria for assessing disorder because they are, in and of themselves, descriptors of pathological functions.” The only exception was obsessive-compulsive personality disorder, where the criteria seemed independent of disordered functioning. In a review of the treatment of personality disorders, Sanislow and McGlashan (12) noted that clinicians regard some personality disorder criteria as symptoms or symptomatic behaviors and as such as legitimate targets of treatment (e.g., stress-related paranoia, suicidal behavior). In contrast, other criteria are reflections of personality traits or style and are considered irrelevant (or resistant) to intervention (e.g., perfectionism, irritability, proclivity to shame). Similarly, Zanarini et al. (13) considered the criteria for borderline personality disorder to be a mélange of acute symptoms, temperamental traits, or amalgams of both.

Although personality disorder criteria are considered heterogeneous and are often criticized because of this feature, taxonomic investigations of schizotypal personality disorder, borderline personality disorder, avoidant personality disorder, and obsessive-compulsive personality disorder highlight the homogeneity of within-category criteria sets. Previous studies have investigated the inter- and intra-consistency of personality disorder criteria cross-sectionally and over time and the stability of the criteria longitudinally.

In an earlier study from our research group, Grilo et al. (14) evaluated cross-sectionally the performance characteristics of the DSM-IV personality disorder criteria for schizotypal personality disorder, borderline personality disorder, avoidant personality disorder, and obsessive-compulsive personality disorder in a clinical sample of 668 adults recruited for the Collaborative Longitudinal Personality Disorders Study (15, 16). The personality disorder criteria sets for all four personality disorders demonstrated convergent validity. The criteria for the individual personality disorders correlated better with each other than with criteria for other personality disorders, i.e., the criteria for all four personality disorders were internally consistent to comparable degrees. Two smaller studies with homogeneous patient study groups (17, 18) also reported findings generally consistent with the baseline Collaborative Longitudinal Personality Disorders Study (14). In a small non-clinical sample evaluated for all DSM-IV personality disorders, however, internal consistency of the criteria sets varied considerably by disorder (19). For the four personality disorders in question, internal consistency of criterion sets was highest for avoidant personality disorder, intermediate for borderline personality disorder and schizotypal personality disorder, and lowest for obsessive-compulsive personality disorder, suggesting more heterogeneity of expression of criteria for the latter in non-treatment-seeking samples.

The temporal coherence of criterion change over 2 years for the four personality disorders investigated in the Collaborative Longitudinal Personality Disorders Study was also evaluated (20). The observed change in each criterion over 2 years was correlated with the observed change in every other criterion over 2 years to determine if there was within-syndrome consistency in the changes. The observed criterion change correlates were consistent within each syndrome (median alpha=0.72 across the four personality disorders) and reasonably specific to that syndrome relative to other disorders. The results supported the validity of these personality disorder criterion sets as representing coherent syndromes.

Two studies of the Collaborative Longitudinal Personality Disorders Study sample have provided information about the longitudinal stability of these criteria. Shea et al. (21) and Grilo et al. (22) reported on the 1-year and 2-year stability, respectively, of schizotypal personality disorder, borderline personality disorder, avoidant personality disorder, and obsessive-compulsive personality disorder as diagnostic categories. Focusing on the 2-year follow-up, significant improvement in the form of diagnostic remission occurred, at rates ranging from 25% in the schizotypal personality disorder sample to 41% in the obsessive-compulsive personality disorder sample. In conjunction with these diagnostic changes, the mean proportion of the criteria met for each of the four personality disorder groups decreased significantly, although a continuous measure of the proportion of the criteria met was significantly correlated. That is, while the number of criteria of each personality disorder decreased over time, the rank-order frequency of the criteria within each personality disorder remained stable. This finding strongly suggests that the criteria constituting specific personality disorders demonstrate a structure as a group that has longitudinal stability.

The generic diagnostic criterion for a personality disorder in DSM-IV is an enduring pattern of inner experience and behavior that is pervasive, inflexible, and of long duration. The study reported here examined the criteria of each of these personality disorders over 2 years to characterize and rank-order them on a hierarchy of prevalence or presence (most to least) and of remission or changeability (least to most). The aim of this study was to identify the criteria that are the most and least enduring for each personality disorder.
Method

Subjects

Study participants were evaluated as part of the Collaborative Longitudinal Personality Disorders Study, a prospective project to examine the longitudinal course of borderline, schizotypal, avoidant, and obsessive-compulsive personality disorders (15). An axis I comparison group meeting the criteria for major depressive disorder but with no personality disorder was also included in the study for contrast. Participants ages 18–45 years were recruited primarily among patients seeking treatment at clinical services affiliated with each of the four recruitment sites in the study; patients with active psychosis, acute substance intoxication or withdrawal, a history of schizophrenia spectrum psychosis, or organicity were excluded. At baseline, the study group comprised 668 participants, 571 of whom met the Diagnostic Interview for Personality Disorders (23) criteria for at least one of the four study personality disorders and 97 of whom displayed major depressive disorder with no personality disorder (for complete demographic, clinical, and comorbidity information, see McGlashan et al. [16]). The current report is based on data for 474 personality disorder patients (83% of the initial study group) for whom complete data through 24 months of follow-up were obtained. No significant baseline differences in diagnostic assignments were observed between retained subjects and those not assessed at the 24-month evaluation ($\chi^2=5.77, df=1, n.s.$).

Procedures

Potential participants were screened by using a self-report questionnaire consisting of items pertaining to the four targeted personality disorders. Eligible participants from whom we obtained informed consent were interviewed in person by experienced and trained interviewers who were monitored and who received regular ongoing supervision. Individual DSM-IV criteria were assessed with the Diagnostic Interview for Personality Disorders (23), a semistructured interview with assessment criteria on a 3-point scale (0=not present, 1=present but of uncertain clinical significance, 2=present and clinically significant). Interrater reliability (based on 84 pairs of raters) kappa coefficients for the four study personality disorders ranged from 0.68 (borderline personality disorder) to 0.73 (avoidant personality disorder); test-retest kappas (based on 52 cases) ranged from 0.63 (schizotypal personality disorder) to 0.74 (obsessive-compulsive personality disorder); median reliability correlations for criteria scores ranged from 0.79 to 0.91 (interrater) and 0.65 to 0.84 (test-retest) (24). Participants were reinterviewed with the Diagnostic Interview for Personality Disorders at 24 months by an interviewer who was blind to all results from the baseline and repeated assessments. The data are presented descriptively.

Results

Criterion Prevalence and Remission

Table 1 details the frequency (percent) of personality disorder criteria with a score of 2 (present and significant) at baseline (column 1) and at 2-year blind follow-up (column 2). Column 3 details the frequency (percent) with which criteria present at baseline (scoring 2) were remitted at 2 years, i.e., had a score of 0 (not present). The values in column 3 do not represent the difference between the values in columns 1 and 2 because column 2 includes criteria that have become newly present and significant between baseline and 2 years. The frequencies are listed by personality disorder diagnostic category, i.e., for patients who met the Diagnostic Interview for Personality Disorders criteria for schizotypal personality disorder (N=85), borderline personality disorder (N=201), avoidant personality disorder (N=266), and obsessive-compulsive personality disorder (N=221). The sum is greater than 474 because many patients had more than one personality disorder.

Table 1 also presents criteria ranked by their presence in each disorder at baseline and 2-year follow-up (most to least) and criteria present at baseline ranked by their rate of remission by 2 years (least to most). The rank ordering highlights the criteria that are both the most prevalent and least changeable over time in each disorder.

Criterion Findings by Disorder

For schizotypal personality disorder, the first six criteria in Table 1 ranked high in frequency at baseline (mean=74%), in contrast to the three observational criteria, which ranked lower (mean=40%). The latter were present considerably less frequently at 2 years (mean=24%), and many that were present at baseline had remitted (mean=46%). In contrast, the (reported) schizotypal personality disorder criteria of paranoid ideation, ideas of reference, odd beliefs, and unusual experiences were among the most prevalent and least changeable criteria.

For borderline personality disorder, all criteria were highly prevalent at baseline. Affective instability, anger, and impulsivity were the most frequent, and identity disturbance, abandonment fears, and self-injury were the least frequent, although still with a frequency of at least 60%. By 2 years the prevalence of the criteria decreased approximately 25%–30%, but the rank ordering of prevalence was exactly the same as at baseline. The rank ordering of criteria that remitted (least to most) was almost the same. For borderline personality disorder, impulsivity, anger, and affective instability were the most frequent and stable criteria, and identity disturbance, abandonment fears, and self-injurious behavior were the least frequent and most changeable.

For avoidant personality disorder all criteria were well represented at baseline (all with frequencies of more than 60%), and they tended to keep the same rank order over time vis à vis prevalence and resistance to remission. Feelings of inadequacy, social ineptness, and a need to be certain of being liked before making social contacts were the most prevalent and stable, and worries about shame and risks of exposure (especially at jobs) were the least prevalent and stable.

For obsessive-compulsive personality disorder the criteria were more variably represented at baseline (31%–83% frequency), but they too tended to retain their rank order of prevalence over time. Rigidity, problems delegating, and perfectionism were the most prevalent and stable criteria. Miserliness was the least represented and most variable.
Discussion

The criteria for schizotypal personality disorder, borderline personality disorder, avoidant personality disorder, and obsessive-compulsive personality disorder, despite limitations in available empirical evidence for their development, have undergone only minor revisions since their introduction in DSM-III. Despite the phenomenological heterogeneity of the DSM-IV personality disorder criteria sets—with criteria representing a variety of traits and symptomatic behaviors and reflecting sometimes normal and sometimes pathological dimensions of personality in clinical samples—these sets demonstrate high internal consistency by disorder both cross-sectionally and over time. The criteria also retain their rank order of prevalence over time within the personality disorder category, despite personality disorder syndromal and criterion improvement (remission).

A key strength of the study was the inclusion of a large number of subjects with clinically significant personality disorders who were assessed with operational criteria by raters trained to reliable standards (24) and followed up by raters blind to prior diagnostic data. The shortcomings were that not all DSM-IV personality disorders were represented and that the results may not generalize to non-

<table>
<thead>
<tr>
<th>Diagnostic Group</th>
<th>Criteria Present at Baseline (Most to Least Frequent)</th>
<th>Criteria Present at 2 Years (Most to Least Frequent)</th>
<th>Criteria Remitted at 2 Years (Least to Most Frequent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizotypal personality disorder patients (N=85)</td>
<td>Paranoid ideation 84</td>
<td>Paranoid ideation 58</td>
<td>Paranoid ideation 20</td>
</tr>
<tr>
<td></td>
<td>Unusual experiences 80</td>
<td>Unusual experiences 54</td>
<td>Ideas of reference 26</td>
</tr>
<tr>
<td></td>
<td>Odd beliefs 76</td>
<td>Odd beliefs 48</td>
<td>Odd beliefs 28</td>
</tr>
<tr>
<td></td>
<td>Ideas of reference 76</td>
<td>Ideas of reference 48</td>
<td>No friends 31</td>
</tr>
<tr>
<td></td>
<td>Social anxiety 72</td>
<td>No friends 38</td>
<td>Unusual experiences 32</td>
</tr>
<tr>
<td></td>
<td>No friends 58</td>
<td>Odd thinking 36</td>
<td>Odd thinking 38</td>
</tr>
<tr>
<td></td>
<td>Odd thinkingb 47</td>
<td>Social anxiety 32</td>
<td>Social anxiety 46</td>
</tr>
<tr>
<td></td>
<td>Odd behaviorb 39</td>
<td>Odd behavior 22</td>
<td>Constricted affect 47</td>
</tr>
<tr>
<td></td>
<td>Constricted affectb 38</td>
<td>Constricted affect 13</td>
<td>Odd behavior 52</td>
</tr>
<tr>
<td>Borderline personality disorder patients (N=201)</td>
<td>Affective instability 95</td>
<td>Affective instability 63</td>
<td>Impulsivity 21</td>
</tr>
<tr>
<td></td>
<td>Anger 87</td>
<td>Anger 57</td>
<td>Anger 22</td>
</tr>
<tr>
<td></td>
<td>Impulsivity 81</td>
<td>Impulsivity 55</td>
<td>Affective instability 23</td>
</tr>
<tr>
<td></td>
<td>Unstable relations 79</td>
<td>Unstable relations 53</td>
<td>Unstable relations 29</td>
</tr>
<tr>
<td></td>
<td>Emptiness 71</td>
<td>Emptiness 45</td>
<td>Stress/paranoia 35</td>
</tr>
<tr>
<td></td>
<td>Stress/paranoia 68</td>
<td>Stress/paranoia 44</td>
<td>Emptiness 39</td>
</tr>
<tr>
<td></td>
<td>Identity disturbance 61</td>
<td>Identity disturbance 35</td>
<td>Identity disturbance 40</td>
</tr>
<tr>
<td></td>
<td>Abandonment fears 60</td>
<td>Abandonment fears 31</td>
<td>Abandonment fears 46</td>
</tr>
<tr>
<td></td>
<td>Self-injury 60</td>
<td>Self-injury 30</td>
<td>Self-injury 46</td>
</tr>
<tr>
<td>Avoidant personality disorder patients (N=266)</td>
<td>Feels inadequate 93</td>
<td>Feels inadequate 62</td>
<td>Socially inept 19</td>
</tr>
<tr>
<td></td>
<td>Socially inept 90</td>
<td>Socially inept 62</td>
<td>Feels inadequate 24</td>
</tr>
<tr>
<td></td>
<td>Preoccupation with rejection 88</td>
<td>Preoccupation with rejection 53</td>
<td>Need to be liked before making social contacts 28</td>
</tr>
<tr>
<td></td>
<td>Need to be liked before making social contacts 82</td>
<td>Need to be liked before making social contacts 51</td>
<td>Making social contacts 34</td>
</tr>
<tr>
<td></td>
<td>Avoids interpersonal jobs 67</td>
<td>No risks, fears embarrassment 44</td>
<td>No risks, fears embarrassment 34</td>
</tr>
<tr>
<td></td>
<td>No risks, fears embarrassment 64</td>
<td>Fears ridicule, shame 38</td>
<td>Fears ridicule, shame 34</td>
</tr>
<tr>
<td></td>
<td>Fears ridicule, shame 62</td>
<td>Avoids interpersonal jobs 31</td>
<td>Avoids interpersonal jobs 35</td>
</tr>
<tr>
<td>Obsessive-compulsive personality disorder patients (N=221)</td>
<td>Problems with delegating 83</td>
<td>Rigidity 52</td>
<td>Rigidity 24</td>
</tr>
<tr>
<td></td>
<td>Rigidity 79</td>
<td>Problems with delegating 51</td>
<td>Pack rat 27</td>
</tr>
<tr>
<td></td>
<td>Perfectionism 79</td>
<td>Perfectionism 44</td>
<td>Problems with delegating 30</td>
</tr>
<tr>
<td></td>
<td>Pack rat 63</td>
<td>Pack rat 41</td>
<td>Problems with delegating 34</td>
</tr>
<tr>
<td></td>
<td>Concern with rules, details, lists 61</td>
<td>Concern with rules, details, lists 32</td>
<td>Perfectionism 35</td>
</tr>
<tr>
<td></td>
<td>Inflexible about morality 58</td>
<td>Workaholic 29</td>
<td>Workaholic 35</td>
</tr>
<tr>
<td></td>
<td>Workaholic 49</td>
<td>Inflexible about morality 27</td>
<td>Concern with rules, details, lists 39</td>
</tr>
<tr>
<td></td>
<td>Misery 31</td>
<td>Misery 10</td>
<td>Inflexible about morality 41</td>
</tr>
</tbody>
</table>

\(^a\) Participants were ages 18–45 years and were recruited from among patients seeking treatment at clinical services affiliated with the sites of the Collaborative Longitudinal Personality Disorders Study.

\(^b\) Observational criterion.
treatment-seeking personality disorder populations. With these strengths and limitations in mind, we present some implications that follow from the data.

The polythetic nature of the DSM-IV criteria for these disorders has often been criticized for its lack of a cohesive, prototypic hierarchy of characteristics and the fact that the system gives equal weight to criteria that may be less central to the personality disorder category they define. Indeed, we found differences among the criteria within each personality disorder—differences in prevalence and stability (or resistance to change) that reflect differences in the nature of the criteria that make up personality disorders. The criteria that are more frequent and enduring over time may reflect elements of personality or personality disorder that are closer to temperament and trait (constitutional proclivities to perceiving and acting/reacting). In contrast, those that are less pervasive and more changeable may be closer to symptomatic behaviors that are stress responsive and habitual (i.e., learned). The former relate more to nature, i.e., genetics and biology; the latter relate more to nurture and learning. The former may be prime targets for biological treatments; the latter, better targets for psychosocial interventions.

Hyman (25) has called for classifying personality disorders on the basis of dimensions that cut across existing categories within axis II and between axis II and axis I. Furthermore, Hyman suggested that the selection of particular dimensions should be based on “empirical factors such as heritability.” Our effort here was an attempt to identify potential core dimensions based on longitudinal prevalence and resistance to change as the parameters of external validity.

Based on these parameters, the criteria to emerge in borderline personality disorder were affective instability, anger, and impulsivity. These criteria reflect what others regard as core trait distortions or endophenotypes of borderline personality disorder, such as affective dysregulation/instability (26–32) or impulsive aggression (26, 32, 33). They reflect two dimensions that emerge recurrently in factor analyses of borderline personality disorder—dysregulated affect and dysregulated behavior (34, 35). They also reflect the time-varying course of the Collaborative Longitudinal Personality Disorders Study borderline personality disorder subjects, with affective dysregulation/instability associated with axis I major depressive disorder and posttraumatic stress disorder (36). It may be that these trait criteria are closer to the core of borderline personality disorder’s biogenetic structures. Furthermore, the less pervasive and more changeable criteria such as self-injury or frantic efforts to avoid abandonment may be seen as secondary or reactive, insofar as such behaviors represent attempts to adapt to, defend against, or cope with pathological affective dysregulation and impulsive aggression (37).

The trait-like criteria that emerged for avoidant personality disorder were regarding oneself as socially inept, feeling inadequate compared to others, and wanting evidence of being liked first before making social contacts. The common theme appears compatible with the internalizing dimension of anxious-misery identified by Kendler et al. (38), a dimension resulting largely from the effects of genetic risk factors. The criteria perhaps reflect the early temperaments of shyness and behavioral inhibition, temperaments that intermittently find symptomatic behavioral expression in a variety of avoidant behaviors (39).

The criteria that emerge as most common and trait-like for schizotypal personality disorder were paranoid ideation, ideas of reference, odd beliefs, and unusual experiences. These criteria probably represent milder variants of the cognitive distortion of reality that is central to the schizophrenia spectrum (40–42). In schizotypal personality disorder this distortion exists in attenuated form and only intermittently becomes expressed behaviorally as oddness or coldness.

Less is known or hypothesized concerning underlying trait dimensions for obsessive-compulsive personality disorder. In fact, our longitudinal criterion data may provide the first clues of the existence and nature of such dimensions. The most prevalent/least changeable obsessive-compulsive personality disorder criteria were rigidity, perfectionism, and problems delegating; these criteria highlight elements of withholding, resistance to change, and the need to control. Do they, perhaps, suggest traits relating to the neurobiology of aggressive control that are intermittently expressed behaviorally as misersliness and/or strict morality?

Our findings carry implications for criterion selection for borderline personality disorder, schizotypal personality disorder, avoidant personality disorder, and obsessive-compulsive personality disorder in DSM-V. Insofar as the concept of stability and resistance to change remains central to the generic definition of axis II, the criteria emerging as most prevalent and least changeable over time are prime candidates for retention. Criteria that are less common and more changeable may require more scrutiny, or they may need to offer other advantages in order to be retained. For example, self-injury is one of the least prevalent and most remitting criteria of borderline personality disorder, yet as a symptomatic behavior it has high visibility and substantial diagnostic efficiency (positive predictive power) cross-sectionally (10, 14), over time (unpublished 2004 study by C. M. Grilo et al.), and across ethnically diverse samples (17). Similarly, the criteria with the highest cross-sectional diagnostic positive predictive power are the symptomatic behaviors such as (observed) odd thinking for schizotypal personality disorder, avoids interpersonal situations for avoidant personality disorder, and concern with rules, details, and lists for obsessive-compulsive personality disorder (14). Clearly, the criteria for these disorders vary in their utility as they do in their source.

Our findings may also shed light on the longitudinal instability of these personality disorders as diagnostic entities (21, 22), that is, the symptomatic behavioral criteria.
“remit” more quickly and more frequently than trait criteria and are largely responsible for dips below the DSM diagnostic threshold for personality disorder. Such criteria may be good markers of disorder (e.g., the high diagnostic efficiency of self-injury for borderline personality disorder) but not good criteria for the assessment of stability of personality disorder pathology.

In conclusion, the DSM-IV criteria for schizotypal personality disorder, borderline personality disorder, avoidant personality disorder, and obsessive-compulsive personality disorder vary in their longitudinal prevalence and stability within disorder. This variation suggests that these DSM-IV personality disorders are hybrids of more stable traits and less stable symptomatic behaviors. The variation also suggests that both sets of criteria are key to defining personality disorders—one set highlighting personality, the other set highlighting disorder.

References

7. Spitzer RL, Endicott J, Gibbon M: Crossing the border into borderline personality and borderline schizophrenia: development of criteria. Arch Gen Psychiatry 1979; 36:17–24

Received March 15, 2004; revision received May 17, 2004; accepted May 20, 2004. From the Collaborative Longitudinal Personality Disorders Study, Department of Psychiatry and Human Behavior, Brown University, Providence, R.I.; the Department of Psychiatry, Columbia University College of Physicians and Surgeons and New York State Psychiatric Institute, New York; the Department of Psychiatry, Harvard Medical School and McLean Hospital, Belmont, Mass.; and the Department of Psychiatry, Yale University School of Medicine and Yale Psychiatric Institute. Address correspondence and reprint requests to Dr. McGlashan, Yale University School of Medicine, 301 Cedar St, New Haven, CT 06519; thomas.mcglashan@yale.edu (e-mail). Supported by NIMH grants MH-50837, MH-50838, MH-50839, MH-50840, and MH-50850, and by Senior Scientist Award MH-01654 to Dr. McGlashan.
Countertransference Phenomena and Personality Pathology in Clinical Practice: An Empirical Investigation

Ephi Betan, Ph.D.
Amy Kegley Heim, Ph.D.
Carolyn Zittel Conklin, Ph.D.
Drew Westen, Ph.D.

Objective: This study provides initial data on the reliability and factor structure of a measure of countertransference processes in clinical practice and examines the relation between these processes and patients’ personality pathology.

Method: A national random sample of 181 psychiatrists and clinical psychologists in North America each completed a battery of instruments on a randomly selected patient in their care, including measures of axis II symptoms and the Countertransference Questionnaire, an instrument designed to assess clinicians’ cognitive, affective, and behavioral responses in interacting with a particular patient.

Results: Factor analysis of the Countertransference Questionnaire yielded eight clinically and conceptually coherent factors that were independent of clinicians’ theoretical orientation: 1) overwhelmed/disorganized, 2) helpless/inadequate, 3) positive, 4) special/overinvolved, 5) sexualized, 6) disengaged, 7) parental/protective, and 8) criticized/mistreated. The eight factors were associated in predictable ways with axis II pathology. An aggregated portrait of countertransference responses with narcissistic personality disorder patients provided a clinically rich, empirically based description that strongly resembled theoretical and clinical accounts.

Conclusions: Countertransference phenomena can be measured in clinically sophisticated and psychometrically sound ways that tap the complexity of clinicians’ reactions toward their patients. Countertransference patterns are systematically related to patients’ personality pathology across therapeutic approaches, suggesting that clinicians, regardless of therapeutic orientation, can make diagnostic and therapeutic use of their own responses to the patient.

Freud first introduced the concept of countertransference in 1910, noting that the patient’s influence on the analyst’s unconscious feelings can interfere with treatment. This early and relatively narrow view of countertransference as an impediment to treatment prevailed in the psychoanalytic literature for several decades. Over time, however, theorists broadened the concept, recognizing that the clinician’s reactions to the patient (conscious and unconscious, emotional and cognitive, intrapsychic and behavioral) may have diagnostic and therapeutic relevance and can, if properly used, facilitate rather than inhibit treatment (1–5).

According to this expanded view, just as the patient’s behaviors with the therapist could provide in vivo insight into his or her repetitive interpersonal patterns and associated thoughts, feelings, and motives, so, too, could the clinician’s responses to the patient provide insight into patterns the patient wittingly or unwittingly evokes from significant others. Klein (6) suggested that the patient may induce the clinician to experience the feelings that the patient is having trouble acknowledging (7) or may draw the clinician into enactments that reflect the patient’s enduring expectations of relationships (8, 9). Sandler (10) introduced the concept of role responsiveness, in which the therapist acts in accordance with a role that is part of a relationship paradigm the patient unconsciously re-creates with the therapist. Wachtel (11, 12) proposed the similar concept of cyclical psychodynamics, by which patients’ fears, wishes, expectations, and behaviors often create self-fulfilling prophecies.

Although the clinical literature on countertransference is rich and rapidly expanding, the corresponding empirical literature is limited (13–17). Research with largely nonclinical samples has provided indirect support for some of these ideas, demonstrating that depressed people tend to elicit criticism from significant others that matches their own self-criticism (18) and that people who are sensitive to rejection tend (through needy, angry, and otherwise distancing behavior) to elicit rejection and hence to confirm and reinforce their internal working models of relationships (19). Giesler and colleagues (20) demonstrated that some of these processes occur in clinical settings as well. A series of analogue studies (21–26) attempted to operationalize the concept of countertransference, defining countertransference responses as therapists’ reactions to patients that are based on the therapists’ unresolved conflict and operationalizing countertransference in terms of avoidant behaviors (e.g., disapproval, silence, ignoring,
narcissistic personality disorder. Najavits and colleagues (27) developed the Ratings of Emotional Attitudes to Patients by Treaters scale, a clinically subtle measure of countertransference designed primarily to study therapists’ response to patients in treatment for substance abuse.

The present study provides initial data on the reliability and factor structure of a clinician-report measure of countertransference processes designed to assess countertransference, broadly defined to include the range of cognitive, affective, and behavioral responses therapists have to their patients. Although the concept of countertransference emerged from psychoanalytic theory and practice, our goal was to devise a measure that could be used by clinicians of any theoretical orientation, so that we could assess the extent to which particular countertransference responses are specific to certain forms of therapy and so that clinicians of any orientation who are trying to get a better diagnostic sense of the patient or a better understanding of what is happening in the therapeutic dyad can make use of the instrument by comparing their own responses to normed psychometric data. Our primary aims were 1) to describe the factor structure and reliability of a broad-band measure of countertransference phenomena and patients’ personality pathology. Thus, our goals were to provide both initial validity data for the measure and a test of clinically derived hypotheses that have never been put to empirical test. In addition, to illustrate the potential clinical and empirical uses of the instrument, we derived a prototype of the “average expectable countertransference response” to patients with narcissistic personality disorder.

Method

Participants

Participants were 181 clinicians who constituted a random national sample of experienced psychiatrists and psychologists from the membership registers of the American Psychiatric Association and American Psychological Association. We requested mailing lists of clinicians with at least 3 years’ postlicensure or postresidency experience who indicated that they performed at least 10 hours per week of direct patient care. As in prior research with this method, psychologists responded at a substantially higher rate than did psychiatrists to the solicitation, allowing us to assess for biases imposed by differential training or response rates. We found no differences between patients described by psychologists and those described by psychiatrists on any variable of interest despite a roughly 3:1 response rate ratio. (Variables of interest included age, sex, race, socioeconomic status, education level, treatment length, and countertransference factor scores [14 t tests, not significant at p<0.01].) We also compared the patients of this sample of clinicians to those of the first 181 clinicians in another sample with whom we used a similar method but paid a substantially higher honorarium and obtained a correspondingly higher response rate (30%) for the psychiatrists; we found no differences between the samples of patients. Together, these data suggest that any potential biases in the tendency to respond had minimal impact on the representativeness of the sample (see also the discussion of limitations in the Discussion section).

Inclusion and Exclusion Criteria

To obtain a cross-section of psychotherapy patients seen in clinical practice, we asked clinicians to describe a nonpsychotic patient at least 18 years old whom they had treated for a minimum of eight sessions (to maximize the likelihood that they would know the patient well enough to provide a reasonably accurate description of the patient). To minimize selection biases, we directed clinicians to consult their calendar to select the last patient they saw during the prior week who met study criteria. Each clinician described only one patient in order to minimize rater-dependent biases. Clinicians received a modest honorarium ($85) for a procedure that took 3–4 hours to complete, with a response rate of approximately 10%.

Procedure

Clinicians could participate either by pen-and-paper forms or on an interactive web site (http://www.psychsystems.net). Web versus paper participants did not differ on any variable studied here (e.g., countertransference factor scores; eight t tests). Clinicians provided no identifying information about the patient (such as name, initials, or social security number) and were instructed to use only information already available to them from their contacts with the patient so that data collection would not compromise patient confidentiality or interfere in any way with ongoing clinical work.

Measures

We employed a number of measures in standardized sequence. We describe those of relevance to the present study here.

Clinical Data Form. The Clinical Data Form (see references 28, 29) assesses a range of variables relevant to demographics, diagnosis, and etiology. Clinicians first provide basic demographic data on themselves, including discipline (psychiatry or psychology), theoretical orientation, employment sites (e.g., private practice, inpatient unit, school), and sex, and then provide data on the patient’s age, sex, race, education level, socioeconomic status, axis I diagnoses, etc. After completing basic demographic and diagnostic questions, clinicians complete ratings of the patient’s adaptive functioning, developmental history, and family history (which will not be described further here).

Axis II diagnosis. To assess axis II disorders, we asked clinicians to rate as present or absent each criterion of each of the DSM-IV axis II diagnoses, randomly ordered. This procedure provides both a categorical diagnosis of each disorder (obtained by applying DSM-IV cutoffs) and a dimensional measure (number of criteria met for each disorder). Our research group and others have successfully used similar measures in a number of investigations (29–31).

Countertransference Questionnaire. The Countertransference Questionnaire (32) is a 79-item clinician-report questionnaire designed to provide a normed, psychometrically valid instrument for assessing countertransference patterns in psychotherapy for both clinical and research purposes. (The instrument can be downloaded at http://www.psychsystems.net/lab.) The items measure a wide range of thoughts, feelings, and behaviors expressed by therapists toward their patients. We derived the 79 items by reviewing the clinical, theoretical, and empirical literature on countertransference and related variables and by soliciting the advice of several experienced clinicians to review the initial item set for comprehensiveness and clarity. We wrote the items in everyday language, without jargon, so that the instrument could be used comparably by clinicians of any theoretical orientation. Items assess a range of responses, from relatively
specific feelings (e.g., “I feel bored in sessions with him/her.”) to complex constructs such as “projective identification” (e.g., “More than with most patients, I feel like I’ve been pulled into things that I didn’t realize until after the session was over.”).

Results

Sample Characteristics

The clinician sample consisted of 141 (77.9%) psychologists and 40 (22.1%) psychiatrists; 58.6% (N=106) of the clinicians were male. The majority saw patients in private practice (N=145, 80.1%), but they also worked in other settings, including hospital (N=57, 31.5%), forensic (N=15, 8.3%), clinic (N=14, 7.7%), or school (N=9, 5.0%) settings. (As might be expected, psychiatrists were more likely to have primary or secondary employment in hospital settings.) The most common self-reported theoretical orientations included psychodynamic (N=73, 40.3%), eclectic (N=55, 30.4%), and cognitive behavioral (N=37, 20.4%).

Reflecting our efforts to obtain a patient sample stratified by sex, about one-half of the patients were male and one-half were female, with an average age of 40.5 years (SD=13.4). The sample was predominantly Caucasian (N=168, 92.8%). Most were middle class (N=102, 56.4%), with 2.8% (N=5) rated as poor, 24.3% (N=44) as working class, and 16.6% (N=30) as upper class. The mean Global Assessment of Functioning Scale score was 58.0 (SD=12.9). Length of treatment averaged 19 months (SD=30.0), with a median of 13 months, indicating that the clinicians knew the patients very well. The most common diagnoses reported by the clinicians were major depressive disorder (N=89, 49.2%), dysthymic disorder (N=68, 37.6%), general anxiety disorder (N=46, 25.4%), and adjustment disorder (N=45, 24.9%).

Factor Structure of the Countertransference Questionnaire

To identify the factor structure of the Countertransference Questionnaire, we first subjected the items to a principal-component analysis using Kaiser’s criteria (eigenvalues >1). We used the scree plot, percentage of variance accounted for, and parallel analysis (33–35) to select the number of factors to rotate. The scree plot indicated a break between eight and nine factors, and parallel analysis indicated eight factors with eigenvalues larger than would be expected by chance (p<0.05 with 100 random data sets). We therefore conducted factor analyses with seven, eight, and nine factors to maximize interpretability.

Several factors emerged across algorithms, rotations, and estimation procedures. We report here the most coherent solution, the eight-factor promax (oblique) solution using maximum likelihood estimation, which we favored a priori because of the characteristics of promax rotations (notably the absence of the assumption of orthogonal factors and the tendency to maximize factor loadings within factors) and maximum likelihood estimation (notably the advantages for use in subsequent confirmatory factor analyses). This solution accounted for 69% of the variance and included factors well marked by at least five items each, suggesting a stable factor structure unlikely to be substantially affected by sample size (36).

Table 1 presents the factor structure. To create factor-based scores for use in this and subsequent studies, we included items loading ≥0.50 for factors 1 and 2, ≥0.40 for factor 3, and ≥0.375 for factors 4–8 to maximize reliability (coefficient alpha). Intercorrelations among the eight factors ranged from –0.16 to 0.58, with a median of 0.30.

Factor 1, overwhelmed/disorganized (coefficient alpha=0.90), was marked by items indicating a desire to avoid or flee the patient and strong negative feelings, including dread, repulsion, and resentment. The items accord with clinical descriptions of countertransference reactions to patients with axis II cluster B disorders, notably borderline personality disorder and narcissistic personality disorder, and with research on disorganized and unresolved attachment patterns (e.g., references 37, 38).

Factor 2, helpless/inadequate (coefficient alpha=0.88), included items describing feelings of inadequacy, incompetence, hopelessness, and anxiety.

Factor 3, positive (coefficient alpha=0.86), was marked by items indicating the experience of a positive working alliance and close connection with the patient.

Factor 4, special/overinvolved (coefficient alpha=0.75), was marked by items describing a sense of the patient as special, relative to other patients, and by items describing “soft signs” of problems in maintaining boundaries, including self-disclosure, ending sessions on time, and feeling guilty, responsible, or overly concerned about the patient.

Factor 5, sexualized (coefficient alpha=0.77), included items describing sexual feelings toward the patient or experiences of sexual tension.

Factor 6, disengaged (coefficient alpha=0.83), included items describing feeling distracted, withdrawn, annoyed, or bored in sessions.

Factor 7, parental/protective (coefficient alpha=0.80), was marked by items describing a wish to protect and nurture the patient in a parental way, above and beyond normal positive feelings toward the patient.

Factor 8, criticized/mistreated (coefficient alpha=0.83), included items describing feelings of being unappreciated, dismissed, or devalued by the patient.

Ruling Out Theoretical Bias as a Rival Hypothesis

This factor structure is conceptually coherent and clinically recognizable. However, an important question is the extent to which its coherence simply reflects the theoretical beliefs of participating clinicians, particularly given that 40% of clinicians in the sample reported a psychodynamic orientation. To evaluate this possibility, we conducted a second factor analysis, this time eliminating all clinicians who reported a psychoanalytic or psychody-
### TABLE 1. Factor Structure of the Countertransference Questionnaire

<table>
<thead>
<tr>
<th>Factor and Item</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1: overwhelmed/disorganized</strong></td>
<td></td>
</tr>
<tr>
<td>I feel resentful working with him/her</td>
<td>0.72</td>
</tr>
<tr>
<td>I wish I had never taken him/her on as a patient</td>
<td>0.71</td>
</tr>
<tr>
<td>When checking my phone messages, I feel anxiety or dread that there will be one from him/her</td>
<td>0.69</td>
</tr>
<tr>
<td>S/he frightens me</td>
<td>0.67</td>
</tr>
<tr>
<td>I feel used or manipulated by him/her</td>
<td>0.62</td>
</tr>
<tr>
<td>I return his/her phone calls less promptly than I do with my other patients</td>
<td>0.61</td>
</tr>
<tr>
<td>I call him/her between sessions more than my other patients</td>
<td>0.60</td>
</tr>
<tr>
<td>I think or fantasize about ending the treatment</td>
<td>0.59</td>
</tr>
<tr>
<td>I feel mistreated or abused by him/her</td>
<td>0.59</td>
</tr>
<tr>
<td>I feel pushed to set very firm limits with him/her</td>
<td>0.56</td>
</tr>
<tr>
<td>I feel angry at him/her</td>
<td>0.52</td>
</tr>
<tr>
<td>I feel repulsed by him/her</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Factor 2: helpless/inadequate</strong></td>
<td></td>
</tr>
<tr>
<td>I feel I am failing to help him/her or I worry that I won’t be able to help him/her</td>
<td>0.84</td>
</tr>
<tr>
<td>I feel incompetent or inadequate working with him/her</td>
<td>0.80</td>
</tr>
<tr>
<td>I feel hopeless working with him/her</td>
<td>0.78</td>
</tr>
<tr>
<td>I think s/he might do better with another therapist or in a different kind of therapy</td>
<td>0.67</td>
</tr>
<tr>
<td>I feel overwhelmed by his/her needs</td>
<td>0.62</td>
</tr>
<tr>
<td>I feel less successful helping him/her than other patients</td>
<td>0.62</td>
</tr>
<tr>
<td>I feel anxious working with him/her</td>
<td>0.61</td>
</tr>
<tr>
<td>I feel confused in sessions with him/her</td>
<td>0.52</td>
</tr>
<tr>
<td><strong>Factor 3: positive</strong></td>
<td></td>
</tr>
<tr>
<td>I look forward to sessions with him/her</td>
<td>0.69</td>
</tr>
<tr>
<td>S/he is one of my favorite patients</td>
<td>0.67</td>
</tr>
<tr>
<td>I like him/her very much</td>
<td>0.67</td>
</tr>
<tr>
<td>I find it exciting working with him/her</td>
<td>0.58</td>
</tr>
<tr>
<td>I am very hopeful about the gains s/he is making or will likely make in treatment</td>
<td>0.52</td>
</tr>
<tr>
<td>I have trouble relating to the feelings s/he expresses</td>
<td>0.48</td>
</tr>
<tr>
<td>If s/he were not my patient, I could imagine being friends with him/her</td>
<td>0.44</td>
</tr>
<tr>
<td>I feel like I understand him/her</td>
<td>0.43</td>
</tr>
<tr>
<td>I feel pleased or satisfied after sessions with him/her</td>
<td>0.43</td>
</tr>
<tr>
<td><strong>Factor 4: special/overinvolved</strong></td>
<td></td>
</tr>
<tr>
<td>I disclose my feelings with him/her more than with other patients</td>
<td>0.64</td>
</tr>
<tr>
<td>I self-disclose more about my personal life with him/her than with my other patients</td>
<td>0.64</td>
</tr>
<tr>
<td>I do things for him/her, or go the extra mile for him/her, in way that I don’t do for other patients</td>
<td>0.52</td>
</tr>
<tr>
<td>I feel guilty when s/he is distress or deteriorates, as if I must be somehow responsible</td>
<td>0.39</td>
</tr>
<tr>
<td>I end sessions overtime with him/her more than with my other patients</td>
<td>0.39</td>
</tr>
<tr>
<td><strong>Factor 5: sexualized</strong></td>
<td></td>
</tr>
<tr>
<td>I find myself being flirtatious with him/her</td>
<td>0.99</td>
</tr>
<tr>
<td>I feel sexually attracted to him/her</td>
<td>0.89</td>
</tr>
<tr>
<td>I feel sexual tension in the room</td>
<td>0.78</td>
</tr>
<tr>
<td>I tell him/her I love him/her</td>
<td>0.62</td>
</tr>
<tr>
<td><strong>Factor 6: disengaged</strong></td>
<td></td>
</tr>
<tr>
<td>I feel bored in sessions with him/her</td>
<td>0.82</td>
</tr>
<tr>
<td>My mind often wanders to things other than what s/he is talking about</td>
<td>0.72</td>
</tr>
<tr>
<td>I don’t feel fully engaged in sessions with him/her</td>
<td>0.53</td>
</tr>
<tr>
<td>I lose my temper with him/her</td>
<td>0.46</td>
</tr>
<tr>
<td>I watch the clock with him/her more than with my other patients</td>
<td>0.46</td>
</tr>
<tr>
<td>I feel annoyed in sessions with him/her</td>
<td>0.42</td>
</tr>
<tr>
<td><strong>Factor 7: parental/protective</strong></td>
<td></td>
</tr>
<tr>
<td>I feel like I want to protect him/her</td>
<td>0.69</td>
</tr>
<tr>
<td>I feel nurturant toward him/her</td>
<td>0.68</td>
</tr>
<tr>
<td>I have warm, almost parental feelings toward him/her</td>
<td>0.67</td>
</tr>
<tr>
<td>I wish I could give him/her what others never could</td>
<td>0.53</td>
</tr>
<tr>
<td>I feel angry at people in his/her life</td>
<td>0.45</td>
</tr>
<tr>
<td><strong>Factor 8: criticized/mistreated</strong></td>
<td></td>
</tr>
<tr>
<td>I feel unappreciated by him/her</td>
<td>0.75</td>
</tr>
<tr>
<td>I feel criticized by him/her</td>
<td>0.63</td>
</tr>
<tr>
<td>I feel dismissed or devalued</td>
<td>0.60</td>
</tr>
<tr>
<td>I feel am “walking on eggshells” around him/her, afraid that if I say the wrong thing s/he will explode, fall apart, or walk out</td>
<td>0.56</td>
</tr>
<tr>
<td>I have to stop myself from saying or doing something aggressive or critical</td>
<td>0.44</td>
</tr>
</tbody>
</table>

---

* Based on data provided by a national random sample of psychiatrists and clinical psychologists (N=181) who were asked to describe a randomly selected nonpsychotic psychotherapy patient at least 18 years old whom they had treated for at least eight sessions.

* Items with high loadings on each factor are listed. As is standard in factor-analytic studies, a substantial number of items (N=25) did not load strongly on any single factor and hence are not listed here. Whether the unlisted items will be deleted from the instrument or reworked in the future will depend on 1) replication of the factor structure in a second sample by using confirmatory factor analysis and 2) whether those items prove useful in composite portraits of different disorders.
COUNTERTRANSFERENCE

The primary findings are reported in Table 2. As predicted, cluster A showed a significant association with the criticized/mistreated factor, although it was not correlated with the disengaged factor. The data strongly supported the associations for cluster B, except for the special/overinvolved factor. The cluster B (dramatic/erratic) disorders showed an additional (unpredicted) association with the disengaged factor and a negative correlation with positive countertransference. The data supported the hypothesis for cluster C.

In secondary analyses, we followed up on some of these patterns (particularly those that were contrary to our expectations) with hypotheses specific to particular disorders (rather than clusters) using partial correlation analysis, holding constant the other nine disorders in each analysis. We hypothesized that borderline personality disorder would show the expected association with the special/overinvolved factor. This hypothesis was supported (partial r=0.23, df=170, p=0.002). We also hypothesized that narcissistic personality disorder would account for the correlation between cluster B disorders and the disengaged factor. In fact, whereas the other cluster B disorders showed no significant associations with therapist disengagement, narcissistic personality disorder was associated with this factor (partial r=0.30, df=170, p<0.001).

Countertransference Responses to Narcissistic Personality Disorder Patients

To illustrate the uses of the Countertransference Questionnaire in clinical practice and to examine the extent to which it can be used to create empirical prototypes of common countertransference patterns in specific types of pathology, we created a composite description of countertransference patterns in the treatments of patients who met the DSM-IV criteria for narcissistic personality disorder. We standardized (z-scored) the items across patients and then averaged the item scores from patients meeting the DSM-IV criteria for narcissistic personality disorder assessed from the axis II checklist. By standardizing items (setting means to 0) before aggregating, we reduced the salience of items that were descriptive of all patients in the sample (e.g., positive feelings) but not specific to patients with narcissistic personality disorder.

Table 3 presents the items most and least descriptive of therapists’ descriptions of countertransference responses to patients with narcissistic personality disorder (N=13). The composite description is remarkably similar to theoretical and clinical accounts (e.g., references 39–42). Clinicians reported feeling anger, resentment, and dread in working with narcissistic personality disorder patients; feeling devalued and criticized by the patient; and finding themselves distracted, avoidant, and wishing to terminate the treatment.

As with the factor analysis, to see whether this portrait of countertransference responses to narcissistic personality disorder patients could be accounted for by clinicians’ theoretical preconceptions, we created a second composite description excluding the data for the three clinicians reporting a psychoanalytic or psychodynamic orientation (N=10). The items and means were virtually identical, suggesting once again that the data were not theory dependent.
Discussion

The results point to several conclusions. First, we identified eight countertransference dimensions that were robust across extraction methods and rotations: 1) overwhelmed/disorganized, 2) helpless/inadequate, 3) positive, 4) special/overinvolved, 5) sexualized, 6) disengaged, 7) parental/protective, and 8) criticized/mistreated. These dimensions are clinically and theoretically coherent, representing diverse reactions clinicians may have toward patients that likely reflect a combination of the therapist’s own dynamics, responses evoked by the patient, and the interaction of patient and therapist.

The factor structure offers a complex portrait of countertransference processes that is substantially more nuanced than global distinctions between positive and negative countertransference. For example, factor analysis identified an overwhelmed/disorganized pattern of countertransference response, characteristic of clinicians’ response to primarily axis II cluster B patients, which bears substantial similarities to descriptions of disorganized attachment in young children and “unresolved” attachment patterns in adults (37, 38). This factor was distinct from other forms of negative countertransference, notably feeling helpless and inadequate, disengaged, and mistreated by the patient. Similarly, we identified three forms of connection with the patient that have elements of closeness—sexualized, special/overinvolved, and parental/protective—that represent both positive feelings as well as potential countertransference snare. This complexity is consistent with clinical observation. What this study suggests, however, is a way of transcending some of the limitations inherent in clinical theories derived from case studies, in which a single clinician attempts to classify countertransference experiences or constellations based on his or her own experience with a limited number of patients. By using an instrument that provides a “common language” (see reference 43) for describing a subtle clinical phenomenon, we can essentially pool the knowledge of dozens of clinical observers, identifying latent constructs (varieties of countertransference experience) that reflect patterns that individual observers themselves may not have recognized.

Second, although every clinician and every therapeutic dyad is distinct, the significant correlations between the countertransference factors and personality disorder symptoms suggest that countertransference responses occur in coherent and predictable patterns. To put it another way, patients not only elicit idiosyncratic responses from particular clinicians (based on the clinician’s history and the interaction of the patient’s and the clinician’s dynamics) but also elicit what we might call average expectable countertransference responses, which likely resemble responses by other significant people in the patient’s life. The associations between countertransference patterns and personality disorder characteristics support the broad view of countertransference reactions as useful in the diagnostic understanding of the patient’s dynamics, particularly those involving repetitive interpersonal patterns. To the extent that patients sharing diagnostic features on axis II have similar ways of thinking, feeling, and behaving interpersonally, one would expect them to evoke similar reactions from others, including therapists, and this appears to be the case.

Third, data from clinicians of different theoretical orientations showed similar patterns vis-à-vis patients with particular kinds of pathology, suggesting that the results

<table>
<thead>
<tr>
<th>TABLE 3. Standardized Countertransference Questionnaire Items Most and Least Descriptive of Therapists’ Response to Patients Meeting the DSM-IV Criteria for Narcissistic Personality Disorder (N=13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Countertransference Item</td>
</tr>
<tr>
<td>I feel annoyed in sessions with him/her</td>
</tr>
<tr>
<td>I feel used or manipulated by him/her</td>
</tr>
<tr>
<td>I lose my temper with him/her</td>
</tr>
<tr>
<td>I feel mistreated or abused by him/her</td>
</tr>
<tr>
<td>I feel resentful working with him/her</td>
</tr>
<tr>
<td>I talk about him/her with my spouse or significant other more than my other patients</td>
</tr>
<tr>
<td>I feel I am “walking on eggshells” around him/her, afraid that if I say the wrong thing s/he will explode, fall apart, or walk out</td>
</tr>
<tr>
<td>When checking my phone messages, I feel anxiety or dread that there will be one from him/her</td>
</tr>
<tr>
<td>I feel unappreciated by him/her</td>
</tr>
<tr>
<td>At times I dislike him/her</td>
</tr>
<tr>
<td>I have to stop myself from saying or doing something aggressive or critical</td>
</tr>
<tr>
<td>My mind often wanders to things other than what s/he is talking about</td>
</tr>
<tr>
<td>I feel criticized by him/her</td>
</tr>
<tr>
<td>I feel angry at him/her</td>
</tr>
<tr>
<td>I watch the clock with him/her more than with my other patients</td>
</tr>
<tr>
<td>I get enraged at him/her</td>
</tr>
<tr>
<td>I dread sessions with him/her</td>
</tr>
<tr>
<td>I feel dismissed or devalued</td>
</tr>
<tr>
<td>I feel like I’m being mean or cruel to him/her</td>
</tr>
<tr>
<td>I feel hopeless working with him/her</td>
</tr>
<tr>
<td>I feel sexual tension in the room</td>
</tr>
<tr>
<td>I think or fantasize about ending the treatment</td>
</tr>
<tr>
<td>I feel bored in sessions with him/her</td>
</tr>
<tr>
<td>I wish I had never taken him/her on as a patient</td>
</tr>
<tr>
<td>I feel interchangeable—that I could be anyone to him/her</td>
</tr>
<tr>
<td>I don’t feel fully engaged in sessions with him/her</td>
</tr>
<tr>
<td>I return his/her phone calls less promptly than I do with my other patients</td>
</tr>
<tr>
<td>I feel like my hands have been tied or that I have been put in an impossible bind</td>
</tr>
<tr>
<td>I feel envious of, or competitive with him/her</td>
</tr>
<tr>
<td>I feel frustrated in sessions with him/her</td>
</tr>
<tr>
<td>Least descriptive</td>
</tr>
<tr>
<td>I like him/her very much</td>
</tr>
<tr>
<td>I feel compassion for him/her</td>
</tr>
<tr>
<td>I am very hopeful about the gains s/he is making or will likely make in treatment</td>
</tr>
<tr>
<td>I look forward to sessions with him/her</td>
</tr>
<tr>
<td>S/he is one of my favorite patients</td>
</tr>
</tbody>
</table>

*Means are in standard deviation units and describe the number of standard deviations the average narcissistic personality disorder patient in the sample is rated above the sample mean on each item.*
are not artifacts of clinicians’ theoretical preconceptions. What is striking about this finding is that coherent patterns of countertransference response emerge in treatments regardless of whether the clinician even “believes” in the concept of countertransference responses or has been trained to attend to them. Put another way, although the concept of countertransference emerged from psychoanalytic observation, the data suggest that clinicians of all theoretical persuasions should attend to and, where possible, make use of information provided in the context of the therapeutic relationship, including their own responses to the patient.

Finally, the empirical portrait of countertransference responses toward patients with narcissistic personality disorder points to the way researchers can use this measure to create empirical prototypes of subtle countertransference constellations with patients presenting with specific types of personality disturbance. In principle, with a large enough sample, one could empirically map the terrain of countertransference patterns in response to multiple forms of personality pathology. One could also identify distinct constellations within diagnoses (e.g., different kinds of narcissistic patients) or to patients who share certain experiences (e.g., survivors of childhood sexual trauma) that may occur across treatments, at different points in therapy, or at different points in a single therapy hour. In working with survivors of childhood sexual abuse, for example, clinicians often face the opposite danger of pushing too much or too early for the patient to remember—and potentially recapitulating the patient’s subjective experience of unwanted penetration, abuse, or lack of boundaries—versus avoiding discussion of traumatic events in intimate detail for fear of traumatizing the patient—and potentially recapitulating the patient’s experience of unacknowledged but shared secrets or the inability or unwillingness of a caregiver who knew about the abuse to talk about it. Identification of such patterns as common constellations in the treatment of abuse survivors could be very useful in teaching clinicians about potential countertransference dangers inherent in working with abuse survivors in a way that is both clinically sensitive and empirically grounded.

**Limitations**

This study has three primary limitations. First, although clinicians are the most obvious informants to report on their own countertransference responses, the countertransference measure we used shares the inherent limits of self-report measures, such as defensive biases and failure to recognize processes that an outside observer might identify. Thus, it would have been useful to have ratings of therapy process by an independent observer (perhaps based on audiotaped sessions) to identify patterns of clinicians’ behavior that would likely converge with clinicians’ self-reports in some ways and diverge from them in others. A related concern is that clinicians provided all the data and hence that their responses on one questionnaire may not be independent of their responses on others (e.g., that their diagnosis of narcissistic personality disorder may not have been independent of their observations of the patient’s behavior in the room with them). This limitation is common to virtually all studies in psychiatry, in which a single observer (usually the patient) provides data on one measure that are then correlated with data from another measure completed by the same informant (either by self-report questionnaire or structured interview). It would have been preferable to collect diagnostic data independently of clinicians’ reports of their countertransference responses, and future research should clearly do so.

Several factors, however, mitigate the concern that the results primarily reflect clinicians’ biases or preconceptions. First, as noted earlier, clinicians of widely different theoretical perspectives and with widely different training (M.D. training versus Ph.D. training) produced highly similar data. If, for example, cognitive behavior clinicians share a theory of countertransference with psychoanalytic therapists, we are unaware of such a theory. Indeed, the similarity across theoretical orientations is one of the most interesting findings of this study, suggesting that patients’ interpersonal patterns are quite robust in the face of different technical styles. Second, because we used dimensional rather than categorical measures of axis II pathology, clinicians’ responses on the countertransference measure were not likely to be influenced by their beliefs about whether the patient had one personality disorder or another. Third, previous research suggests that clinicians tend to make highly reliable and valid judgments if their observations and inferences are quantified using psychometric instruments such as the ones used in this study. For example, correlations between treating clinicians’ and independent interviewers’ assessments of a range of variables, such as measures of personality pathology and adaptive functioning, tend to be large, typically >0.50 (44–46). Empirically, clinicians’ theoretical orientation predicts little variance in descriptions of clinical phenomena when clinicians were asked to describe a specific patient rather than their beliefs or theories of psychopathology (47, 48). Nevertheless, future research using this measure should assess the association between countertransference phenomena and patients’ personality pathology by using data about the patient provided by other observers.

The second limitation was clinicians’ response rate to our request for participation (approximately 10%, for a study described as requiring 3–4 hours of work for a token honorarium). Three factors, however, limit the likelihood that the results reflect response rate biases. First and foremost, it is hard to imagine a response rate hypothesis that could explain the pattern of results. By virtue of their willingness to donate 3–4 hours of their time, the clinicians who participated in the study may have been characterized by greater interest in research, altruism, financial dis-
tress, or a host of other factors, compared with colleagues who did not participate, but it is difficult to see how any of these variables could have produced the obtained findings. Second, because we solicited data from clinicians in multiple practice settings across North America and provided them with a method of randomly selecting a patient within their practice, we were able to obtain a broad cross-section of patients. (Clinicians who agreed to participate were unaware that countertransference was one of the constructs we intended to study, so we were not selecting clinicians with a particular interest in or knowledge of this domain.) Third, as noted earlier, psychologists’ response rate was almost three times the rate of psychiatrists, yet the two sets of informants provided similar data, suggesting that neither training nor response rate was responsible for the findings. Finally, all studies have selection biases, some avoidable and others less so. An alternative method would have been to sample clinicians at a single hospital or institution, but this strategy would likely have resulted in less generalizable results. It seems unlikely that the sample in our study was less representative of the population of clinicians or patients in psychotherapy than samples in typical studies of psychotherapy or psychopathology, which tend to rely on patient populations from a single site. Nevertheless, the data clearly require replication and extension with new groups of patients, and we hope this article will generate interest in the measure for future research.

A third potential objection is sample size, given the possibility of some instability of factor structure with a ratio of cases to items of <3:1. However, recent thinking about factor analysis, based on data from Monte Carlo simulations and other studies, suggests that factor solutions stabilize with far fewer cases than previously believed (typically by 100 cases), as long as the factors are well marked by a sufficient number of items with loadings above 0.40 or 0.50 (as they were here), and that conventional case-to-item ratios do not take into consideration a range of variables that qualifies them in one direction or the other (see, e.g., references 36, 49). Clearly the next step in this research, however, is a larger-N replication study using confirmatory factor analysis and external ratings of variables such as personality disorder diagnosis and treatment outcome independent of the clinician’s reports.

Implications

The Countertransference Questionnaire represents an effort to develop a readily administered measure that reflects shared clinical wisdom in its item content and statistical “wisdom” in its factor structure. This measure is germane to future research on countertransference phenomena, as well as to practice, allowing clinicians to clarify the diagnostic relevance and utility of their reactions by comparing their own responses to normed psychometric data. A broadband measure of countertransference processes with known correlates such as this can turn clinicians’ experiences into quantifiable dimensions that capture interpersonal patterns that emerge in sessions, allowing clinicians who normally attend to countertransference phenomena to hone and systematize their self-reflections and providing clinicians whose theoretical orientations do not emphasize such processes a language and method with which to capture information about the patient and the treatment process that may be diagnostically and therapeutically significant.

Received March 17, 2004; revision received July 13, 2004; accepted Aug. 18, 2004. From Georgia School of Professional Psychology; Emory University Psychoanalytic Institute, Atlanta; Hoover & Associates, Orland Park, Ill.; Cambridge Hospital/Harvard Medical School, Cambridge, Mass.; Harvard Medical School, Boston; and the Departments of Psychology and Psychiatry and Behavioral Sciences, Emory University. Address correspondence and reprint requests to Dr. Betan, Georgia School of Professional Psychology, 980 Hammond Dr., Suite 100, Atlanta, GA 30328; ebetan@argosy.edu (e-mail); or Dr. Westen, Departments of Psychology and Psychiatry and Behavioral Sciences, Emory University, 532 Kilgo Circle, Atlanta, GA 30322; dwesten@emory.edu (e-mail).

Supported in part by NIMH grants MH-62377 and MH-62378 to Dr. Westen and by a grant from the Fund for Psychoanalytic Research of the American Psychoanalytic Association to Dr. Heim.

The authors thank the 181 clinicians who contributed data to this study.

References

COUNTERTRANSFERENCE

17. Dube J, Normandin L: The mental activities of trainee therapists of children and adolescents: the impact of personal psychotherapy on the listening process. Psychotherapy 1999; 33: 216–228
Dissociation, Childhood Interpersonal Trauma, and Family Functioning in Patients With Somatization Disorder

Richard J. Brown, Ph.D., Clin.Psy.D.
Anette Schrag, M.D., Ph.D.
Michael R. Trimble, M.D.

Objective: The goals of this study were to determine 1) the occurrence of various dissociative phenomena in patients with somatization disorder, 2) the occurrence of six different types of childhood interpersonal trauma in these patients, and 3) the nature of these patients’ early family environment.

Method: Twenty-two patients with somatization disorder and 19 medical comparison subjects completed the Structured Clinical Interview for DSM-IV Dissociative Disorders, the Childhood Trauma Interview, and the Family Functioning Scale.

Results: The somatization disorder patients reported significantly higher level of dissociative amnesia than the comparison subjects. The two groups reported similar levels of depersonalization, derealization, identity confusion, and identity alteration. Somatization disorder patients reported significantly greater childhood emotional abuse and more severe forms of physical abuse, relative to the comparison subjects, with chronic emotional abuse being the best predictor of unexplained symptoms. Childhood sexual abuse, separation/loss, and witnessing violence were equally common in the two groups. The somatization disorder group reported significantly more family conflict and less family cohesion.

Conclusions: Only some types of dissociation are more severe in patients with somatization disorder, relative to medical comparison subjects. Many patients with somatization disorder are raised in an emotionally cold, distant, and unsupportive family environment characterized by chronic emotional and physical abuse. Sexual abuse is not a necessary prerequisite for the disorder.

Physicians frequently encounter patients with a history of symptoms that cannot be explained by organic factors (1). In cases where such symptoms cannot be attributed to anxiety, depression, or hypochondriasis, a diagnosis of somatoform disorder is made. Somatization disorder, characterized by a history of at least eight unexplained symptoms in four or more bodily systems, represents the extreme end of a continuum of somatoform severity. Such patients describe high levels of distress, functional disability, and health resource utilization (2, 3). Many are resistant to psychological treatment, and the prognosis of the condition is poor (4).

Early attempts to describe the pathogenesis of such “medically unexplained” symptoms assumed that these conditions were caused by the separation, or “dissociation,” of distressing material from conscious awareness (5, 6). This process of dissociation is thought to be triggered by traumatic events, often occurring in childhood. The concept of dissociation has since been resurrected in a number of theories, and it has become the focus of one of the dominant models in this area (e.g., references 7–10).

Several different psychological symptoms are thought to reflect a dissociative process, including psychogenic amnesia, depersonalization-derealization, and identity alteration. A number of studies have shown that such symptoms are particularly common in patients with somatoform disorders, apparently supporting the dissociation model (11–15). Conversely, many patients presenting with dissociative disorders also report large numbers of unexplained symptoms (16–19). However, not all studies have found a link between dissociation and unexplained illness (20–22). Moreover, one study showed that the relationship between dissociation and unexplained illness disappears in analyses that control for trauma exposure (23). Such inconsistent findings are attributable to confusion about the precise definition of dissociation and to difficulties in operationalizing the phenomenon empirically (24, 25).

Several different forms of dissociation have been identified (e.g., references 5, 6, 7, 10), and an increasing number of phenomena have been associated with the dissociative label (25). Existing measures of dissociation, such as total scores on the Dissociative Experiences Scale (26), tend to conflate these different types of dissociation, making it impossible to assess which, if any, are associated with unexplained illness.

A number of studies have assessed the prediction that somatoform illness is associated with childhood trauma. Somatoform illness has been linked to both childhood...
physical (11, 21, 22, 27–30) and sexual abuse (11, 22, 23, 27–30). However, only a minority of studies (23, 30) have investigated patients with multiple unexplained symptoms. There is also limited evidence to suggest that childhood emotional neglect (22) and abuse (23, 28, 29) are associated with unexplained illness. However, studies addressing emotional abuse have either used nonstandard measures (23) or identified the occurrence of abuse using a single question (27, 28), which is unlikely to provide an accurate measure of trauma prevalence. Accordingly, very little is known about the nature or frequency of emotional abuse in patients with unexplained symptoms. Similarly, existing research has not addressed whether experiences such as neglect and parental loss are also associated with the occurrence of somatoform illness.

Research in this area has other limitations. Many studies, for example, have used dichotomous codes for physical or sexual abuse (i.e., present versus absent), without reporting information about frequency, severity, or duration. This approach prevents analyses that can assess the relationship between trauma magnitude and subsequent psychopathology. In addition, most studies have not considered the fact that childhood trauma is often confounded with other pathogenic family factors, such as boundary confusion, rigid behavioral control, and poor adaptability (31). It may be that these factors play a more important role in the development of somatoform illness than childhood trauma per se.

The present study investigated the occurrence of dissociation and trauma in patients with somatization disorder, extending the literature in several ways. First, we assessed for dissociation using a structured clinical interview that provides information about various types of dissociative phenomena. This step allowed us to address the hypothesis that only some forms of dissociation are associated with somatoform illness (24). Second, we used a standardized semistructured interview that addresses not only childhood physical and sexual abuse but also four other domains of childhood interpersonal trauma (separation/lack, neglect, emotional abuse, and the witnessing of violence) that have rarely been investigated previously. This measure provides information concerning the severity, frequency, and duration of trauma, allowing for an assessment of the relationship between trauma magnitude and symptom patterns. Finally, we included a measure of childhood family functioning to assess the possibility that general family factors mediate the relationship between somatoform illness and childhood trauma.

Method

Participants

Twenty-eight patients with a documented history of multiple unexplained symptoms agreed to participate. All had received in- or outpatient treatment at a specialist neurological hospital and were identified through this center. Twelve were identified retrospectively through analysis of discharge summaries from the hospital neuropsychiatry ward from the preceding 3 years. Ten were identified prospectively after admission to the neuropsychiatry ward and had been identified as having multiple unexplained symptoms during this admission. A further six participants were recruited as part of another study but were included after their histories of multiple unexplained symptoms were uncovered.

Diagnoses of somatization disorder were made according to the DSM-IV criteria. Participants had to have a history of multiple unexplained physical symptoms beginning before age 30 years, including at least four pain symptoms, two gastrointestinal symptoms, one sexual-reproductive symptom, and one neurological symptom. Symptoms had to have precipitated treatment seeking or caused impairment in social/occupational functioning. Symptom histories were obtained by means of an open interview about the participant’s current and past medical history, as well as a structured interview about previous physical symptoms and disorders as part of the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (32). Where possible, general practice and hospital records were obtained (with the patient’s consent) and studied. A neurologist (A.S.) with psychiatric experience conducted the interviews and analyzed all records. Symptoms were classified as either explained or unexplained by organic factors, on the basis of results from appropriate investigations and/or clinical diagnosis. Minor medical symptoms, such as minor injuries, pharyngitis, otitis media, or a single episode of urinary tract infection or gastroenteritis, were excluded.

The comparison group initially consisted of 22 patients with dystonia of established organic origin. These subjects were consecutively recruited from a neurological botulinum toxin clinic and had diagnoses of classical cervical dystonia or generalized dystonia caused by the DYT-1 mutation, a basal ganglia lesion, or anoxic birth injury. All of the comparison participants underwent the same evaluation as the somatization disorder group. It was later found that one of the comparison subjects met the criteria for somatization disorder, and data for that subject were excluded. SCAN data from two comparison subjects and one patient with somatization disorder were lost because of equipment failure. Three participants with unexplained symptoms did not meet the full criteria for somatization disorder, and their data were excluded. Two somatization disorder patients did not complete the dependent measures for logistical reasons.

The final study group consisted of 22 patients with somatization disorder (20 women, two men) with an average age of 40.9 years (SD=9.5) and 19 comparison participants (13 women, six men) with an average age of 47.4 years (SD=14.6). The two groups did not differ with respect to age (t=1.67, df=30, p=0.11) or sex (χ²=3.28, df=1, p=0.07). General practitioner records were obtainable for 14 patients with somatization disorder (73.7%) and 14 comparison subjects (63.6%); hospital records were available for all participants. The mean number of unexplained symptoms in the somatization disorder group was 31.3 (SD=10.6), compared to 3.7 in the comparison group (SD=2.6).

By using SCAN results, the following comorbid DSM-IV disorders were identified in the somatization disorder group: generalized anxiety disorder (N=10), panic attacks (N=9), simple phobia (N=1), agoraphobia (N=3), social phobia (N=1), obsessive-compulsive disorder (N=1), posttraumatic stress disorder (N=3), mood disorder (N=13), and eating disorder (N=1). The following comorbid disorders were identified in the comparison group: generalized anxiety disorder (N=4), anxiety disorder not otherwise specified (N=3), depression not otherwise specified (N=3), and alcohol abuse disorder (N=1). Seven somatization disorder patients and 11 comparison subjects did not meet the criteria for any other clinical disorders.

After complete description of the study, written informed consent was obtained from all participants.
Materials

Dissociation. Dissociation was measured with the Structured Clinical Interview for DSM-IV Dissociative Disorders (SCID-D) (33). The SCID-D is a semistructured interview that assesses past and current episodes of clinical dissociation in five domains: amnesia, depersonalization, derealization, identity confusion, and identity alteration. Each section comprises a series of probe questions pertaining to different examples of dissociation within that domain. Positive responses are followed up by questions concerning the nature and frequency of each experience. In each domain, the participant’s reports are rated on a 4-point scale of severity (absent, mild, moderate, severe), on the basis of a detailed manual. The scale has good reliability and validity (33).

Childhood trauma. The Childhood Trauma Interview (34) is a brief semistructured interview assessing childhood interpersonal trauma in six domains: separation/loss, neglect, emotional abuse, physical abuse, witnessing violence, and sexual abuse. Each section comprises a series of probe questions designed to elicit experiences relevant to that domain. Positive responses to probe questions are followed up with questions concerning the nature, frequency, and duration of the experience and the perpetrators involved. This process is repeated until all perpetrators and forms of trauma in each domain are identified. The scale has good psychometric properties and convergent validity (34).

The severity and frequency of childhood interpersonal trauma are rated separately on scales ranging from 1 to 6, on the basis of a detailed manual. Total exposure scores for each trauma domain are calculated by summing the product of the severity, frequency, and duration scores across all perpetrators within each category. The separation/loss section addresses interruptions in early attachments, including those caused by parental absence related to separation, divorce, illness, death, incarceration, etc., or the absence of the child from the family home. The neglect section asks about early material deprivations (e.g., of food, clothing, shelter, medical care) and levels of childhood supervision. The emotional abuse section inquires about experiences of being shouted at, insulted, criticized, threatened, ignored, humiliated, or scapegoated. The physical abuse section asks about experiences of being hit, beaten, kicked, burnt, suffocated, cut, shot, or locked up. The witnessing violence section includes questions on violence witnessed both in and outside the home. The sexual abuse section asks about both contact and noncontact childhood sexual experiences. Consenting experiences with similar-age peers were not included.

Family functioning. The Family Functioning Scale (35) is a 75-item self-report questionnaire assessing family functioning in 15 domains: cohesion, expressiveness, conflict, intellectual-cultural orientation, active-recreational orientation, religious emphasis, organization, family sociability, external locus of control, family idealization, disengagement, democratic family style, laissez-faire family style, authoritarian family style, and enmeshment. It has satisfactory psychometric properties and discriminant validity (35). Scale items consist of statements concerning family life; participants are asked to rate on a 4-point Likert scale how true each statement is for their own family. All participants were asked to complete the scale in relation to their childhood family setting.

Procedure

All participants were interviewed by the second author about their past medical history and were administered the SCAN. After a break, a research psychologist (R.B.) administered the SCID-D and Childhood Trauma Interview. Participants were asked to complete the family functioning questionnaire at home.

Results

Dissociation

Table 1 presents descriptive statistics for the SCID-D and Childhood Trauma Interview. Amnesia scores were significantly higher for the somatization disorder group (U=134.5, p=0.03), with 11 patients (50%) reporting severe dissociative amnesia, compared to three comparison subjects (15.8%). The groups did not differ on any of the other SCID-D dimensions; median scores of 1 on each of the other dimensions for both groups indicated that most participants did not report other dissociative symptoms.

Childhood Trauma

All but three participants (one patient with somatization disorder, two comparison subjects) had experienced moderate to severe trauma (severity score of 3 or above on the Childhood Trauma Interview) of at least one kind. The somatization disorder group reported more severe forms of physical abuse (U=123.0, p<0.04). Regarding total exposure scores, the somatization participants scored significantly higher than the comparison subjects in the domain of emotional abuse (U=104.5, p=0.009); the somatization subjects had greater exposure to physical abuse than the comparison subjects, but the difference only approached significance (U=130.5, p=0.06). The comparison group had significantly higher total exposure scores for neglect (U=133.5, p=0.05). The somatization disorder group also reported more emotional abuse perpetrators (U=123.5, p<0.04) and a longer duration of emotional abuse (U=110.0, p<0.02).

Data Analysis

One participant from each group failed to complete the Family Functioning Scale, one somatization disorder participant refused to complete the Childhood Trauma Interview, and one somatization disorder participant was unable to complete the witnessing violence section of the Childhood Trauma Interview. In addition, one participant with somatization disorder failed to complete approximately half of the items on the Family Functioning Scale. Data for these participants were excluded from all analyses involving these variables. For the remaining participants, 0.97% of the comparison group responses had missing values, as did 1.16% of the somatization disorder group responses. With the exception of Family Functioning Scale items, all missing values were replaced with the mean of each individual’s scores on the remaining items from the relevant subscale.

The majority of the SCID-D and Childhood Trauma Interview variables had L-shaped distributions. Accordingly, medians and interquartile ranges are reported for these variables. For all inferential analyses of these variables, nonparametric statistics (Mann-Whitney test, Spearman’s rank-order correlations) were used. Parametric tests (t tests) were used for comparative analyses involving the Family Functioning Scale. The nonindependence of the study variables meant that statistical correction for multiple comparisons using the Bonferroni or Holm (36) adjustments was inappropriate in this context. To minimize the likelihood of type I errors, all statistical analyses were two-tailed.
Family Functioning

Table 2 presents descriptive statistics for the Family Functioning Scale. The somatization participants scored significantly lower than the comparison subjects on cohesion (t=2.44, df=27.7, p<0.03) and significantly higher on conflict (t=2.35, df=36, p<0.03).

Correlations

Total number of unexplained symptoms correlated significantly with cohesion scores on the Family Functioning Scale (r=−0.57, df=20, p=0.01), as well as with emotional abuse (r=0.70, df=21, p<0.0001) and physical abuse (r=0.48, df=21, p<0.03) scores. The correlation between number of unexplained symptoms and SCID-D amnesia domain only approached significance (r=0.39, df=22, p=0.07). Emotional abuse exposure correlated significantly with SCID-D amnesia domain (r=0.44, df=21, p<0.05).

Discussion

The findings of this study are only partly consistent with previous research and theory suggesting a link between unexplained symptoms and dissociation (5–19). Dissociative amnesia was frequently reported by the somatization disorder participants and was significantly more common in this group, relative to the medical comparison subjects. However, there were no differences in the level of depersonalization, derealization, identity confusion, and identity alteration reported by the two groups.

These findings suggest that only certain dissociative phenomena are associated with unexplained illness (24). One interpretation is that unexplained neurological symptoms share a common etiological mechanism with dissociative amnesia (8, 24), which is different from the mechanism operating in depersonalization-derealization (24). It is noteworthy that only dissociative amnesia has separa-
tion of material from conscious awareness as a cardinal feature, a concept that is central to early psychodynamic definitions of dissociation (5, 6). It may be that these definitions of dissociation are more relevant to understanding unexplained illness than the generic definition adopted by DSM-IV, which combines phenomena such as amnesia, depersonalization, and identity confusion within a common category irrespective of whether the same mechanism is operating in each case. Explicit definition of the types of dissociation being investigated is clearly needed in future studies in this area, and measures that enable different types of dissociative phenomena to be looked at separately should be included. For this reason, the use of total scores on the Dissociative Experiences Scale (26) as a unitary measure of dissociation should be proscribed.

Although the somatization disorder group had been exposed to more severe forms of physical abuse and had witnessed more extreme forms of violence, relative to the comparison subjects, the two groups did not differ significantly in total exposure to physical abuse, contrary to previous research (11, 21, 22, 27). However, the somatization group reported a significantly higher level of family conflict and a significantly lower level of family cohesion. Evidently, many patients with somatization disorder are raised in an environment characterized by frequent arguments, emotional distance, and poor support, consistent with the high levels of physical and emotional abuse reported by this group. The pattern and magnitude of the intercorrelations between these variables suggest that the relationship between emotional abuse and unexplained symptoms in patients with somatization disorder cannot be attributed solely to their exposure to a broadly pathogenic family environment (see reference 30).

One obvious limitation of this study is its reliance on retrospective trauma reports without corroborating evidence. It is possible, for example, that the relatively high levels of trauma reported by the somatization disorder group are due to overreporting. Previous research suggested that overreporting of early trauma is rare, however (37). It is also possible that dissociative amnesia led to underreporting of trauma in the somatization disorder group, which may explain why sexual abuse was rarely reported by these individuals. Nevertheless, these patients did not seem to be amnesic for the chronic physical and emotional abuse that they experienced. This finding raises important questions about the nature of dissociation in patients with unexplained symptoms. According to Janet (5) and Meares (10), any traumatic event leading to the development of unexplained symptoms should be unavailable for recall because of dissociation. If this were true, the traumatic events reported by the somatization disorder subjects could not have prompted the development of unexplained symptoms through dissociative processes. According to Breuer and Freud (6), however, dissociation may simply limit access to the affect associated with traumatic events and not to the memories of the events themselves. Our findings are clearly more consistent with this account of dissociation. Ultimately, methods that consider the dissociative process itself, rather than the antecedents or consequences of this process, are needed to determine the exact nature of dissociation in somatoform illness.

Another limitation of the study is its use of a neurological rather than a psychiatric comparison group. Because of the characteristics of the comparison group, the findings provide no information on whether the levels of emotional and physical abuse reported by the somatization disorder patients are higher than those in patients with other psychiatric conditions. However, comorbid psychopathology was commonly reported by the comparison participants, who were similar to the somatization disorder participants in other illness variables, making them an

### Table 2. Family Functioning Scale Scores of Patients With Somatization Disorder and Comparison Subjects With Dystonia

<table>
<thead>
<tr>
<th>Family Functioning Scale Item</th>
<th>Patients With Somatization Disorder (N=21)</th>
<th>Dystonia Comparison Subjects (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohesion</td>
<td>12.1 ± 2.0</td>
<td>13.3 ± 1.0</td>
</tr>
<tr>
<td>Expressiveness</td>
<td>12.3 ± 1.7</td>
<td>12.4 ± 1.5</td>
</tr>
<tr>
<td>Conflict</td>
<td>10.9 ± 1.8</td>
<td>9.6 ± 1.5</td>
</tr>
<tr>
<td>Intellectual-cultural</td>
<td>12.9 ± 1.9</td>
<td>12.7 ± 1.5</td>
</tr>
<tr>
<td>Active-recreational</td>
<td>12.1 ± 1.5</td>
<td>12.2 ± 1.7</td>
</tr>
<tr>
<td>Religious emphasis</td>
<td>10.7 ± 2.3</td>
<td>11.5 ± 1.9</td>
</tr>
<tr>
<td>Organization</td>
<td>13.4 ± 1.6</td>
<td>13.2 ± 1.5</td>
</tr>
<tr>
<td>Family sociability</td>
<td>12.6 ± 2.7</td>
<td>13.3 ± 2.7</td>
</tr>
<tr>
<td>External locus of control</td>
<td>12.4 ± 2.5</td>
<td>12.9 ± 1.5</td>
</tr>
<tr>
<td>Family idealization</td>
<td>12.5 ± 2.5</td>
<td>13.4 ± 2.1</td>
</tr>
<tr>
<td>Disengagement</td>
<td>12.9 ± 2.3</td>
<td>12.2 ± 1.7</td>
</tr>
<tr>
<td>Democratic family style</td>
<td>11.3 ± 1.6</td>
<td>10.6 ± 1.7</td>
</tr>
<tr>
<td>Laissez-faire family style</td>
<td>10.9 ± 2.1</td>
<td>10.1 ± 1.4</td>
</tr>
<tr>
<td>Authoritarian family style</td>
<td>11.9 ± 1.9</td>
<td>11.4 ± 1.8</td>
</tr>
<tr>
<td>Emmeshment</td>
<td>10.9 ± 3.4</td>
<td>10.3 ± 2.9</td>
</tr>
</tbody>
</table>

a Significant difference between groups (t=2.44, df=27.7, p<0.03).

b Significant difference between groups (t=2.35, df=36, p<0.03).
appropriate comparison group. Moreover, because this study included a medical comparison group, it provides information about the value of dissociation and childhood trauma as aids to the differential diagnosis of patients with unexplained symptoms. Nevertheless, attempts at replication with larger groups of subjects and with other comparison groups are necessary to establish the validity and generalizability of the study findings.

Taken together, these findings suggest that chronic emotional abuse might be the most important setting condition for the development of somatization disorder; other severe forms of interpersonal trauma may occur in this context, but they perhaps play a lesser role in the development of pathology. The study indicates that many people with somatization disorder are exposed to an early environment that is emotionally cold, harsh, and characterized by frequent criticism, insults, rejection, and physical punishment. This environment also appears to be linked to the development of dissociative amnesia, a common concomitant of unexplained illness. However, it is apparent that other dissociative phenomena are no more common in somatization disorder than in comparison medical populations. Similarly, sexual abuse does not appear to be a necessary prerequisite for the development of multiple unexplained symptoms.

Received Feb. 25, 2003; revision received March 3, 2004; accepted June 14, 2004. From the Department of Clinical Neurology, Institute of Neurology, London; and the Academic Division of Clinical Psychology, University of Manchester. Address correspondence and reprint requests to Dr. Brown, Academic Division of Clinical Psychology, University of Manchester, Rawnsley Building, Manchester Royal Infirmary, Oxford Rd., Manchester, M13 9WL, UK; richard.james.brown@manchester.ac.uk (e-mail).

Supported by the Raymond Way Fund of the Institute of Neurology. The authors thank Maria Ron and E.S. Krishnamoorthy for their comments.

References

18. van der Kolk BA, Pelcovitz D, Roth S, Mandel FS, McFarlane A, Herman JL: Dissociation, somatization, and affect dysregulation: the complexity of adaptation to trauma. Am J Psychiatry 1996; 153(July) festchrift suppl;83–93
Temperament and Character Profiles and the Dopamine D4 Receptor Gene in ADHD

Deborah E. Lynn, M.D.
Gitta Lubke, Ph.D.
May Yang, M.S.
James T. McCracken, M.D.
James J. McGough, M.D.
Janeen Ishii
Sandra K. Loo, Ph.D.
Stanley F. Nelson, M.D., Ph.D.
Susan L. Smalley, Ph.D.

Objective: This study was designed to investigate the link among attention deficit hyperactivity disorder (ADHD) in adults, novelty-seeking temperament, and the 48-base pair (bp) dopamine D4 receptor (DRD4) gene variant.

Method: This study drew from a larger molecular genetic study of ADHD in which the ascertainment criterion was having an affected sibling pair with ADHD. Parents (N=171) from 96 families provided data. Of the 171 parents, 56 (33%) had a lifetime history of ADHD, with 28 (50%) continuing to meet DSM-IV criteria (i.e., “persistent” ADHD). Latent variable modeling was used to test whether the DRD4 gene variant or Temperament and Character Inventory factors could predict ADHD.

Results: Using latent variable modeling, the authors were able to confirm the first-order factor structure of the Temperament and Character Inventory. Furthermore, novelty seeking predicted ADHD lifetime diagnosis (R²=26%), while the DRD4 gene variant independently predicted ADHD (R²=5%) but not novelty seeking.

Conclusions: In this unique sample of parents from multiply affected ADHD families, novelty seeking and the 48-bp DRD4 variant were associated with a lifetime history of ADHD. However, the association between novelty seeking and ADHD does not appear to be due to variation in the 48-bp DRD4 variant.

For the complex disorder of attention deficit hyperactivity disorder (ADHD), one challenge is to elucidate the individual differences in developmental pathways leading to various ADHD outcomes. Many researchers are trying to understand how genetic liability interacts with environment in the development of ADHD and clinical variability (1–5). Among our first attempts to illuminate the pathway from genotype to phenotype is an investigation of the relationship between genes, temperament, and ADHD outcome.

The dopamine D4 receptor (DRD4) gene has consistently been widely investigated in the quest to discover the genetic underpinnings of ADHD, with the 48-base pair (bp) variant at the focus of most studies. Complementing our own literature review (6–21) in 2001, Farone et al. (22) conducted a meta-analysis of 14 case/control and family-based association studies that investigated the 7-repeat allele of the DRD4 48-bp variant and ADHD. In the meta-analysis, they found significant associations for both family-based and case/control studies (odds ratios of 1.4 and 1.9, respectively). Although to our knowledge no comparable meta-analysis exists for DRD4 and novelty seeking, there are some 23 studies that have investigated the 48-bp repeat variant and novelty seeking, with 52% of these showing positive results (24–44).

The results of these studies suggest a possible role of DRD4 in both ADHD and novelty seeking, although the estimated effect size is very small and results remain tenuous because of the large number of studies reporting negative findings. Three factors are likely to contribute significantly to the variability in findings, including 1) the polygenic nature of ADHD and temperament and the difficulty in replicating minor gene effects under polygenic inheritance (45), 2) clinical variability across studies, and 3) the possibility of false positive in light of numerous analyses and nominal p values used to determine significance for many candidate gene investigations. To what extent DRD4 plays a role in the genetic liability to ADHD, novelty seeking, or both thus warrants further investigation. The possible common genetic underpinnings of novelty seeking and ADHD on the basis of individual associations with the same “risk” DNA variant would suggest that they would overlap to some extent at a phenotypic level. However, there is a paucity of research on the relationship of novelty seeking and ADHD and, to our knowledge, no investigation of DRD4 in the context of a joint analysis of the two phenotypes in any one group of subjects.

Temperament has been conceptualized as a genetically influenced building block of personality; it has been shown to be highly heritable and relatively stable across the lifespan (46). Personality has been suggested to represent a developmental outcome of the interplay of environmental factors with temperament over time (47–49). The role of extremes of temperament in the development of psychopathology is well documented in mood and anxiety disor-
found that hyperactivity-impulsivity and oppositional conscientiousness and high neuroticism. In addition, they counted for by his or her own temperament profile (55). In a subsequent study of adults, Nigg et al. (56) found a relationship between novelty seeking and harm avoidance. The investigators used the Tridimensional Personality Questionnaire, a precursor to the more comprehensive temperament scales of the Temperament and Character Inventory. The investigators also found that these ADHD adults had high rates of comorbid depressive disorder, antisocial personality disorder, and alcohol and drug abuse/dependence (47% had a current axis I anxiety or depressive disorder, and 37% had a comorbid conduct disorder or mood disorder as children).

In related studies, Nigg and Hinshaw and their colleagues (54) found that ADHD subjects (N=78) scored significantly higher than normal subjects on the temperament scales of novelty seeking and harm avoidance. The investigators used the Temperament and Character Inventory questionnaire and are included in the present study.

Method

Cohort

Subjects were 171 parents (87 women and 84 men; mean age=43.4 years [SD=6.2]) from 96 families identified from an ongoing molecular genetic study of ADHD in which families had been ascertained through an ADHD-affected sibling pair. After complete description of the study to the subjects, written informed consent was obtained. Among parents, 56 (33%) had a lifetime history of ADHD, meaning they qualified for a definite or probable lifetime diagnosis of ADHD regardless of current ADHD status. Of these 56 parents, 50% (N=28) showed “persistent” ADHD (i.e., continued to meet criteria into adulthood for either definite or probable ADHD). Among the parents, 87% were Caucasian, 5% were Hispanic, 5% were African American, and 3% were of mixed ancestry. The socioeconomic status (per Hollingshead 1957 rankings) breakdown of the subjects was as follows: class I=11%, class II=33%, class III=19%, class IV=11%, class V=9%, class VI=2%, class VII=15%. Most (84%) were married, 14% were separated or divorced, 2% had remarried, and 1% never married. About half (51.4%) of the individual parents had completed at least 4 years of college. For a full description of the sample assessment procedures, see Smalley et al. (58). The 96 families represent a subset of families who were administered the Temperament and Character Inventory (52), a measure added several years into the data collection protocol. After it was added, all subsequent subjects answered the Temperament and Character Inventory questionnaire and are included in the present study.

Measures

After providing written informed consent approved by the UCLA Institutional Review Board, subjects were assessed with the Schedule for Affective Disorders and Schizophrenia—Lifetime Version Modified for the Study of Anxiety Disorders (63). As a supplement to assess ADHD and conduct disorder, the behavioral disorders section of the Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version (64) was also administered. Interrater reliabilities for ADHD diagnoses were obtained in a subset of 26 interviews rated from tapes with weighted kappas of 0.96, 0.95, and 0.95. Diagnosing ADHD in adult samples has been controversial. For this reason, our diagnostic criteria also included a spousal report when possible (161 of the 171 parents [94%] had spousal reports). Otherwise, adults were asked to recall as best as possible their early behavior as well as report their current symptoms.

The Temperament and Character Inventory (52) is a measure designed to assess differences between individuals in seven basic dimensions of temperament and character. The instrument measures four temperament indices (novelty seeking, harm avoidance,
TABLE 1. Temperament and Character Profiles of 171 Parents From 96 Families With an ADHD-Affected Sibling Pair, by ADHD History, and a Population Comparison Group

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Parents From Families With an ADHD-Affected Sibling Pair</th>
<th>Population Comparison Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No History of ADHD (N=115)</td>
<td>History of ADHD (N=56)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>44 ± 6.5</td>
<td>42 ± 5.4</td>
</tr>
<tr>
<td>Temperament and Character Inventory item</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novelty seeking</td>
<td>8.1 ± 3.8</td>
<td>10.4a</td>
</tr>
<tr>
<td>Harm avoidance</td>
<td>8.2 ± 4.7</td>
<td>9.6 ± 4.9</td>
</tr>
<tr>
<td>Self-directedness</td>
<td>18.9 ± 5.1</td>
<td>17.5 ± 5.2</td>
</tr>
<tr>
<td>Cooperativeness</td>
<td>21.6 ± 3.1</td>
<td>20.7 ± 3.9</td>
</tr>
<tr>
<td>Self-transcendence</td>
<td>5.6 ± 3.8</td>
<td>7.6b</td>
</tr>
</tbody>
</table>

a Significantly different from score of parents with no history of ADHD (z=3.17, two-tailed p=0.002).
b Significantly different from score of parents with no history of ADHD (z=3.59, two-tailed p=0.0003).

Statistical Analysis

The exploratory analysis serves two goals. First, it allows a determination of the number of factors needed to explain the common variance of the observed subscales. We tested whether the Temperament and Character Inventory actually has a seven-factor structure as seen in other normative and clinical samples (50). Second, it provides estimates of the factor loadings of each observed subscale on each of the factors, such that we can evaluate whether the Temperament and Character Inventory subscales cluster in the expected way (i.e., all novelty seeking subscales load on a single factor but not on other factors) (65–67). On the basis of the resulting pattern of factor loadings, a confirmatory model was developed and fit to the data. The confirmatory factor model was then used to test the role of temperament and DRD4 on lifetime ADHD diagnosis as well as ADHD symptom variability.

Results

As shown in Table 1, parents with a history of ADHD had significantly higher scores for novelty seeking and self-transcendence than did the unaffected parents. Overall the parents of ADHD-affected sibling pairs had scores similar to population comparison subjects.

Results of our exploratory factor analyses of the measurement model of the Temperament and Character Inventory are shown in Table 2.

In exploratory analyses, we varied the number of latent factors from three to eight. In all models, factors were allowed to correlate. The percent of variance (R²) pertaining to the persistence subscale was below 0.15, and the reward dependence subscales did not cluster on a single factor in any of the models but loaded on several factors simultaneously. Therefore, given the lack of reliability and interpretability for these subscales, they were omitted from further analyses. The resulting five-factor model (model I) provided an acceptable fit to the data. In the context of the five-factor solution, the self-directedness subscale loaded on two separate factors, so the four-factor solution (model II) corresponded better with the known Temperament and Character Inventory structure. Under the four-factor model, the novelty seeking, cooperativeness, and self-transcendence subscales all load on separate factors, while the self-directedness and harm avoidance subscales load on a single bipolar factor, that is, high harm avoidance subscale in the parents. All analyses were carried out using Mplus, version 2.13 (65, 66).

Genotyping

Blood samples were collected from each family member, and DNA was isolated using the Puregene Kit following the manufacturer's recommendations (Gentra Systems, Minneapolis). A polymorphic region in the DRD4 gene consisting of a variable number of 48-bp repeats was scored by polymerase chain reaction (PCR) with primer flanking the repeat sequence. PCR amplification was performed in 12.5-μl reactions containing 60 ng genomic DNA, 10% DMSO, 400 μM DNTPs, 0.8 μM each primer; 10 mM KCl, 20 mM Tris-HCl (pH 8.8), 10 mM (NH₄)₂SO₄, 0.1% Triton X-100, and one unit Vent DNA polymerase (New England Biolabs, Beverly, Mass.). The primers used were D4-42 5′-GCG ACT ACG TGG TCT ACT CG-3′ and D4-42 5′-ACG AGC TCT ATG GCC TTAG-3′ (Op-eron, Alameda, Calif.). Using the MJ Research PTC-100 thermal cycler, DNA was denatured at 98°C for 7 minutes, followed by 32 cycles of 94°C (1 minute), 54°C (30 seconds), 72°C (2 minutes), and final extension at 72°C for 7 minutes. Final PCR products were electrophoresed in a 2.5% Nu Sieve agarose gel in 1× TBE buffer for 2.5 hours at 100 volts. The gels were ethidium bromide stained for 30 minutes and destained in dH₂O for 1 hour. Alleles were determined by comparison of bands to known molecular weight standards.

Genotypes were coded as 0 or 1, reflecting the absence or presence, respectively, of the putative "risk" allele (the 7-repeat variant). The total number of genotyped subjects was 127 because 44 samples failed to yield sufficient DNA for genotyping after inclusion in other genetic investigations. The genotyping could not be done on these 44 samples because the dilution samples were degraded at the time of this analysis. The mean scores on Temperament and Character Inventory scales and the distribution of ADHD status for these 44 parents unavailable for the molecular analysis were no different from those available for study (data not shown).

In the present study group, the genotype was coded as 0 for 65 parents (67%) and 1 for 42 parents (33%). The homozygous and heterozygous groupings of the 7-repeat allele were pooled because of the small sample size (62%). The homozygous group (7/7) was pooled with the heterozygous group (7/8 and 7/9) to increase the number of subjects in the group, resulting in one group of 89 parents (47% of the study sample) and 1 for 42 parents (33%). The homozygous and heterozygous groupings of the 7-repeat allele were pooled because of the small sample size (62%).
To retain the four-factor solution but allow for the bipolar factor, we defined in the confirmatory model a factor for self-directedness and one for harm avoidance (i.e., the bipolar factor is modeled as two factors (Table 2, Model III). Overall, we were largely able to confirm the factor structure of the Temperament and Character Inventory, particularly for novelty seeking, the prime factor of interest.

The confirmatory factor analysis measurement model (model III) was specified as the model under which we would perform other analyses with respect to ADHD. This model is depicted in Figure 1.

Squares represent the observed variables (the Temperament and Character Inventory subscales) whereas circles indicate the unobserved factors. The arrows represent the factor loadings of observed variables on the unobserved factors. The fit of the confirmatory model is acceptable according to commonly used cutoff criteria (e.g., root mean square error of approximation [RMSEA] <0.06, standardized root mean square residual [SRMR] <0.08, weighted root mean square residual [WRMR] <0.9) (67).

Figure 1 depicts the factor loading estimates. Note that for each factor, one subscale is fixed to 1.0 in order to properly scale the latent factor. In this model, all Temperament and Character Inventory factors are specified to be correlated. The resulting estimated correlations between the novelty seeking, harm avoidance, self-directedness, cooperativeness, and self-transcendence factors are generally low, with the exception of harm avoidance and self-directedness as expected because of their bipolarity being modeled as two correlated factors. We were unable to establish a clear secondary factor structure, featuring two higher-order factors, temperament and character, as proposed by Cloninger et al. (52).

Using this confirmatory factor model (model III), we tested whether factors underlying temperament and character contribute to lifetime ADHD diagnosis or symptom variability. As shown in model IV (Table 2), the Temperament and Character Inventory factors explained 49% of the variance in ADHD diagnosis. We tested the role of novelty seeking relative to other latent constructs of temperament and character on ADHD diagnosis (model V). A comparison of models IV and V revealed that novelty seeking was the greatest contributor to ADHD diagnosis, accounting for 26% of variance in the ADHD phenotype. Since novelty seeking contributed to over half of the variance in ADHD and no other Temperament and Character Inventory factor contributed to such a large extent, we selected Model V as the most parsimonious model for subsequent analyses. No other remaining Temperament and Character Inventory factor contributed individually in a significant manner, as reflected by the overall fit of the model with only the novelty seeking factor included.

In addition to testing the role of novelty seeking on ADHD diagnostic status, we examined the role of novelty seeking (and other temperament scales) on individual symptoms of inattention and hyperactivity-impulsivity by using the sum of inattentive items and the sum of hyperactive-impulsive items generated from the psychiatric interview, again based on clinical symptoms reported during the childhood period. As can be seen in Table 2, novelty seeking and the character dimension cooperativeness accounted for the majority of variability in symptoms of hyperactivity-impulsivity and inattention, as indicated by a comparison of the restricted model for which only novelty

---

**TABLE 2. Model Fitting Analyses of Temperament and Character Inventory Constructs**

<table>
<thead>
<tr>
<th>Model</th>
<th>Goodness of Fit Indicesa</th>
<th>Variance Explained (R²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploratory factor analyses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five-factor model (model I)</td>
<td>π²: 254.19</td>
<td>df: 158</td>
</tr>
<tr>
<td>Four-factor model (model II)</td>
<td>π²: 304.4</td>
<td>df: 206</td>
</tr>
<tr>
<td>Confirmatory factor analysis (model III)</td>
<td>π²: 257.15</td>
<td>df: 158</td>
</tr>
<tr>
<td>Contribution of temperament and character factors to ADHD diagnosis (model IV)</td>
<td>π²: 99.43&lt;sup&gt;c&lt;/sup&gt;</td>
<td>df: 51</td>
</tr>
<tr>
<td>Contribution of novelty seeking alone to ADHD diagnosis (model V)</td>
<td>π²: 103.45&lt;sup&gt;c&lt;/sup&gt;</td>
<td>df: 51</td>
</tr>
<tr>
<td>Contribution of temperament and character factors to symptoms of inattention and hyperactivity-impulsivity (model VI)</td>
<td>π²: 293.00&lt;sup&gt;c&lt;/sup&gt;</td>
<td>df: 188</td>
</tr>
<tr>
<td>Contribution of novelty seeking to symptoms of inattention and of novelty seeking and cooperativeness to symptoms of hyperactivity-impulsivity (model VII)</td>
<td>π²: 303.4&lt;sup&gt;c&lt;/sup&gt;</td>
<td>df: 195</td>
</tr>
</tbody>
</table>

<sup>a</sup> RMSEA=root mean square error of approximation. SRMR=standardized root mean square residual. WRMR=weighted root mean square residual. CFI=comparative fit index.

<sup>b</sup> SRMR criterion for goodness of fit is <0.08.

<sup>c</sup> Chi-square results from weighted least squares estimation, which is used because ADHD is categorical. It is not comparable to chi-square derived from maximum likelihood theories, and it is not appropriate to compare models using a likelihood ratio test.

<sup>d</sup> WRMS criterion for goodness of fit is <0.90.

<sup>e</sup> Models VI and VII use maximum likelihood, which allows for comparison of models using a likelihood ratio test where twice the likelihood difference is distributed as chi-square with df equaling the difference in df between two models, one of which is a restricted version of the other.
seeking and cooperativeness were included (model VII) to the full model in which novelty seeking, harm avoidance, self-directedness, cooperativeness, and self-transcendence were included (model VI) ($\chi^2_{\text{diff}}=10.4$, df diff=7, $p=0.17$). Under the restricted model, Model VII, novelty seeking and cooperativeness predict hyperactivity-impulsivity with an $R^2$ estimate of 0.28, while inattention is predicted by novelty seeking, but not cooperativeness, with an $R^2$ estimate of 0.16.

We examined the influence of DRD4 on ADHD diagnosis and symptom sums and the symptoms of hyperactivity-impulsivity and inattention in the subset of individuals who were genotyped for the DRD4 48-bp polymorphism. We tested the unique contribution of DRD4 to ADHD by running a restricted model eliminating the path from novelty seeking to ADHD and comparing it to a full model including novelty seeking. We extended Model V (Table 2) to include all Temperament and Character Inventory factors again to evaluate DRD4. Dropping the nonsignificant regressions of DRD4 on harm avoidance, self-directedness, and self-transcendence led to a slight increase in model fit due to the increase in parsimony. The model fit well and is shown in Figure 2.

ADHD is predicted by DRD4 ($z=2.87$, $p=0.004$) yet the contribution of DRD4 to novelty seeking is nonsignificant and in the opposite direction (absence of “risk” allele, higher novelty seeking score) than that of ADHD. Eliminating the path from DRD4 to novelty seeking resulted in a negligible change in model fit as reflected by the RMSEA criteria (0.048 to 0.051) and WRMR (0.803 to 0.821). Furthermore, there was a lack of change in variance explained for ADHD ($R^2=0.28$ in both models) with or without DRD4 predicting novelty seeking in the models. The implication of this is that DRD4 and novelty seeking are working more or less independently to influence ADHD, a topic to be examined more in the Discussion section.

Second, we evaluated the role of DRD4 on specific hyperactive-impulsive and inattentive symptoms by extending Models VI and VII in Table 2. We examined and pre-
dicted hyperactivity-impulsivity, inattention, and all Temperament and Character Inventory factors as a function of DRD4. Again, regressions on harm avoidance, self-transcendence, and self-directedness on DRD4 were approximately zero and dropped from the model. The resulting model had a good fit, as reflected by RMSEA=0.041 and comparative fit index=0.941.

The regression weights of DRD4 on novelty seeking and cooperativeness were similar and in the negative direction (i.e., absence of the “risk” allele, higher novelty seeking and cooperativeness scores) while in the positive direction on the hyperactive-impulsive and inattentive symptom sum scores. DRD4 had nonsignificant influences on hyperactivity-impulsivity and inattention, and the only path reaching significance was that of DRD4 on novelty seeking in this model ($z=-2.0$). Inclusion of DRD4, however, led to an overall modest improved fit of the model relative to one without DRD4 (Model VII, Table 2) as reflected by the change in RMSEA (0.057 to 0.041) and SRMR (0.077 to 0.073). Excluding all paths from DRD4 except that predicting novelty seeking resulted in no significant change in goodness of fit of the model ($\chi^2_{\text{diff}}=6.48$, df$_{\text{diff}}=3$, p=0.09).

**Discussion**

The first-order factor structure described by Cloninger’s Temperament and Character Inventory was evident in our dataset. In general, we saw that the observed variables of each construct indeed measured the latent variable as expected. These results strengthen the reliability of the Temperament and Character Inventory as a measure of temperament.

The current study replicates and confirms the findings of Downey et al. (54) in a sample size twice as large. Specifically, there is a strong role of novelty seeking as a predictive factor of ADHD diagnostic status. While correlation does not imply causality, we modeled novelty seeking and other temperament factors as predictors of ADHD rather than ADHD predicting temperament. This directionality was based on the theoretical and empirical research suggesting that temperamental differences are evident in infancy and early childhood (45). Temperament, in general, was found to be a major predictor of lifetime ADHD status in parents of ADHD-affected sibling pairs ($R^2=0.49$), with novelty seeking the primary contributor ($R^2=0.26$). Similar to Downey and colleagues’ findings, novelty seeking and harm avoidance contribute to ADHD status (although harm avoidance was individually not significant). However, novelty seeking was by far the larger contributor to ADHD. One explanation for the discrepancy in the role of harm avoidance in the Downey study and ours may be that our subjects reflect a more homogeneous (and likely genetic) group of ADHD adults because of the fact that they were ascertained through having at least two ADHD-affected children. Perhaps novelty seeking plays a greater role relative to other temperament scales in familial ADHD than in less familial cases. The temperament scale of self-transcendence was actually more strongly associated with ADHD in our sample than harm avoidance, but neither scale individually reached a level of significance to require inclusion in the most parsimonious model.

We also found a significant role of the character dimension cooperativeness in predicting hyperactive-impulsive symptoms but not ADHD status in the parents. Downey et al. did not have a method to assess this dimension because they used an earlier instrument, the Tridimensional Personality Questionnaire, instead of the more current Temperament and Character Inventory. Since cooperativeness reflects items of character maturity (e.g., “I usually accept other people as they are, even when they are very different from me” and “I cannot have any peace of mind if I treat other people unfairly, even if they are unfair to me”), the relationship of cooperativeness in adulthood to hyperactive-impulsive symptoms in childhood may reflect a continuous developmental trajectory. Conversely, the presence of hyperactive-impulsive symptoms in childhood may contribute to poor development of aspects of maturity, e.g., by interfering with social and emotional regulation. Further work exploring the relationship of this aspect of character development with hyperactive-impulsive symptoms in ADHD is needed.

At the diagnostic level, DRD4 plays a minor but significant role. However, the association of ADHD and novelty seeking was not accounted for by the presence of a risk allele at DRD4. When individual symptom scores (hyperactive-impulsive and inattentive) were evaluated rather than ADHD, DRD4 did not contribute to symptom variability but did have a minor effect on novelty seeking. The pattern of weights from DRD4 to ADHD and novelty seeking suggests that the relationship of DRD4 to ADHD is opposite that of DRD4 to novelty seeking, in contrast to the expectation if the putative “risk” DRD4 variant was accounting for their association. These data strongly suggest that the 7-repeat variant of the 48-bp polymorphism at DRD4 is a genetic variant associated in small part with ADHD (<5%) and possibly novelty seeking (<5%), but that it does not account for the strong phenotypic association of the two traits. To our knowledge, two other studies documented a negative relationship of DRD4 to novelty seeking, but most associations of DRD4 and novelty seeking are in the reverse direction. Gelernter and colleagues (29) found this sort of negative association specifically in European American women and substance-dependent African American subjects. Malhotra and colleagues (68) found this negative association specifically in a Finnish population with substance abuse diagnoses. Further work investigating the relationship of DRD4 and novelty seeking in the context of substance dependence or abuse within ADHD samples may clarify the discrepancies observed across studies. DRD4 also contributed a small (and marginally significant) proportion of variance in the character
trait cooperativeness, but again in a negative direction (absence of the allele, higher cooperativeness).

The current investigation of DRD4 as a predictor of ADHD or novelty seeking in parents ascertained from multiply affected ADHD families suggests that the strong association of ADHD and novelty seeking observed in the current sample is not due to the small influence of the DRD4 48-bp 7-repeat variant. A minor role of this genetic variant on ADHD is indicated, but its influence pales in comparison to that of the personality construct novelty seeking (25%) in accounting for “liability” to ADHD. It remains impossible to identify whether novelty seeking increases one’s risk for ADHD or whether the presence of ADHD influences the development of novelty-seeking temperament. Assuming that aspects of temperament emerge before the onset of ADHD symptoms, investigations of high novelty-seeking temperament in infancy may prove useful for identifying “at risk” ADHD children. More important, molecular work investigating the genetics of ADHD may benefit from inclusion of novelty seeking as a potential "endophenotype" given its strong association with ADHD and high heritability (53).

We have furthered the knowledge that parents of children with ADHD, who have a history of ADHD themselves, have significantly high novelty seeking, a temperament factor that may predispose one to ADHD or may be a result of having ADHD. In addition, the parents who have a history of high levels of impulsivity and hyperactivity also have lower scores on the cooperativeness character factor as adults, perhaps suggesting that impulsive behavior leads to reduced development of cooperativeness in adulthood (52). The present finding has important implications for intervention in ADHD through identifying important parent-based issues. Such parental factors may influence the variability in ADHD symptom persistence, comorbidity, and degree of impairment. This becomes an especially poignant question when considering the immense challenges parents of ADHD children face and that success of interventions may vary as a function of parental characteristics. Every clinician has faced the challenges of some families in which the parents as well as the children are affected with ADHD. An illustrative example of this is a case of a 7-year-old boy, brought in by his parents with the complaints of hyperactivity, difficulty with paying attention at school, and difficulty studying at home all causing major problems. As the evaluation got underway, we discovered the mother's lifelong struggle with inattention and disinhibition. This not only impacted her life but caused her to target this boy all throughout his latency years as being “bad” and “unmanageable since birth when his scream was the loudest.” It became clear as the treatment progressed that this boy, the third of four children, was exceptionally challenging for this mother who had struggled with similar symptoms. This complicated the treatment of the boy in many ways. Psychotherapy for the family was helpful to alleviate damage to the boy's self-esteem, and proper medication and psychotherapy for the mother helped alleviate her struggle as well. Although the present study does not review the limited data on specific difficulties presented in these cases of both parents and children having ADHD, our work does suggest that further investigation of this issue is important. In light of major efforts underway to best understand how to intervene and influence the development of ADHD and curtail the development of comorbidity and dysfunction in individuals with ADHD, this study may help guide our efforts of treatment development. Targeting character development or development of coping mechanisms in the parents of ADHD children, as well as in the ADHD children themselves, may serve to improve treatment outcomes. This would mean targeted psychotherapy in some cases with parents in addition to any psychopharmacologic interventions.

Limitations of the present study include the possibility that our results cannot generalize to ADHD adults because parents of ADHD-affected sibling pairs may differ. In addition, ADHD diagnoses in these parents were based on retrospective data, and recall bias may affect findings. Furthermore, high novelty-seeking individuals may differentially report themselves as having ADHD. To buffer against this sort of reporting error, we did use observer ratings, clinical evaluations, and spouse reports to ensure the most reliable diagnosing.

References


http://ajp.psychiatryonline.org

Am J Psychiatry 162:5, May 2005


ADHD, NOVELTY SEEKING, AND DRD4


W. Gordon Frankle, M.D.
Ilise Lombardo, M.D.
Antonia S. New, M.D.
Marianne Goodman, M.D.
Peter S. Talbot, M.D.
Yiyun Huang, Ph.D.
Dah-Ren Hwang, Ph.D.
Mark Slifstein, Ph.D.
Susan Curry, B.A.
Anissa Abi-Dargham, M.D.
Marc Laruelle, M.D.
Larry J. Siever, M.D.

Objective: The serotonin system is believed to play a role in modulating impulsivity and violence. Previous imaging studies have implicated the anterior cingulate and orbitofrontal cortex in impulsive aggression. This study evaluated regional serotonin transporter distribution in the brain of individuals with impulsive aggression by using positron emission tomography (PET) with the serotonin transporter PET radiotracer [11C]McN 5652.

Method: Ten individuals with impulsive aggression and 10 age- and sex-matched healthy comparison subjects underwent [11C]McN 5652 PET. All individuals were medication free at the time of scanning. Regional total distribution volumes were derived by using a one-tissue compartment kinetic model with arterial input function. Outcome measures of serotonin transporter availability included the binding potential and the specific-to-nonspecific partition coefficient ($V_3''$).

Results: Serotonin transporter availability was significantly reduced in the anterior cingulate cortex of individuals with impulsive aggression compared with healthy subjects, as noted by differences in both binding potential (mean=3.1 ml/g [SD=1.9] versus 5.0 ml/g [SD=2.0], respectively) and $V_3''$ (mean=0.15 [SD=0.09] versus 0.26 [SD=0.09]). In other regions examined, serotonin transporter density was nonsignificantly lower in individuals with impulsive aggression compared with healthy subjects.

Conclusions: Pathological impulsive aggressivity might be associated with lower serotonergic innervation in the anterior cingulate cortex, a region that plays an important role in affective regulation.

Reduced activity of the serotonin (5-HT) system has been implicated in impulsive violence and aggression in studies that have used a variety of paradigms, including measurement of CSF serotonin metabolites, hormonal response to serotonergic probes, and imaging metabolic changes with serotonergic agents. In initial studies, reduced CSF concentration of the serotonin metabolite 5-hydroxyindoleacetic acid was demonstrated in individuals with a history of aggression (1, 2). Subsequent studies have linked this finding to impulsive aggression (3). Moreover, a blunted hormonal response to pharmacological manipulation of central serotonin function has been observed in personality disorder patients with impulsive aggression (4, 5).

Studies of brain lesions have pointed to the orbitofrontal cortex and the anterior cingulate gyrus as key areas regulating the generation of aggressive behaviors (6–10). Irritability and angry outbursts have been associated with orbitofrontal cortex damage in neurologic patients (11). Lesions of the medial orbital cortex early in childhood can result in antisocial, disinhibited, aggressive behavior later in life (12). These studies suggest that the orbitofrontal and adjacent medial frontal cortex exert an inhibitory influence on aggressivity. Studies combining positron emission tomography (PET) imaging of regional glucose metabolism with pharmacologic challenges aimed at increasing serotoninergic function also point toward altered 5-HT function in these regions. In healthy subjects, a single dose of the serotonin-releasing agent fenfluramine resulted in an increase in glucose metabolism in the orbitofrontal cortex and medial frontal and cingulate regions. Such an increase was not observed in subjects with impulsive aggression (13). Furthermore, patients with impulsive aggression demonstrated altered metabolic response to the serotonergic agent meta-chlorophenylpiperazine (m-CPP) in the orbitofrontal cortex and the anterior cingulate cortex (14). Together, these findings are consistent with the existence of reduced 5-HT function in the orbitofrontal cortex and anterior cingulate cortex in subjects with impulsive aggression.

In this study, we assess regional brain serotonin innervation in subjects with impulsive aggression by using in vivo imaging of serotonin transporter with PET and [11C](+)-6β-(4-methylthiophenyl)-1,2,3,5,6α,10β-hexahydropyrrolo[2,1-a]isoquinoline (l11C)McN 5652. This radiotracer has been successfully developed to image serotonin transporter density in humans (15–19) and has been used in a number of clinical studies, including studies of patients with mood disorders (20), obsessive-compulsive...
disorder (21), and ecstasy abuse (22, 23). Given that the strongest finding in our [9]CPP study was reduced activation of the cingulate cortex (14), the main hypothesis of this study was that individuals with impulsive aggression would have reduced serotonin transporter density in the anterior cingulate cortex. Serotonin transporter density in the orbitofrontal cortex is too low to be accurately measured with PET and [11C]McN 5652 (18), so this region was not evaluated in this study.

Method

Human Subjects

The study was approved by the institutional review boards of the New York State Psychiatric Institute, Columbia Presbyterian Medical Center, Mount Sinai Hospital, and the Bronx Veterans Affairs Medical Center. Written informed consent was obtained from each subject after explanation of the study procedures. All subjects were free of significant medical problems, had no current or past neurological disorder, had no history of loss of consciousness, were not pregnant or nursing, and had not taken any psychoactive medications in the 3 weeks preceding the PET scan (6 weeks for fluoxetine). Ten patients (five men and five women; mean age=35 years [SD=9, range=18–51]) meeting criteria for intermittent explosive disorder–revised (impulsive aggression) (24) were recruited through advertisements in local newspapers and referrals from outpatient psychiatrists at the Bronx Veterans Affairs Medical Center and Mount Sinai School of Medicine. Patients were considered eligible for the impulsive aggression group if they met research diagnostic criteria for intermittent explosive disorder–revised (25) and the DSM-IV borderline personality disorder “impulsiveness” criterion or the borderline personality disorder “self-damaging” criterion according to a “Module for Intermittent Explosive Disorder-Revised” (prepared by Coccoro et al., personal communication). The Structured Interview for DSM-IV Personality Disorders (26) was used for personality disorder diagnoses. Patients with a history of schizophrenia or other psychotic disorder according to the Structured Clinical Interview for DSM-IV Axis I Disorders (27) were excluded. All subjects with a history of alcohol/drug dependence or substance abuse that had been active in the preceding 6 months were also excluded from the study. Ten healthy comparison subjects (five men and five women; mean age=34 years [SD=8, range=24–49]) with no current or past DSM-IV axis I psychiatric disorder were recruited through advertisements in local newspapers.

The absence of pregnancy and medical and neurological abnormalities was confirmed by a review of the patient’s history and systems, physical examination, routine blood tests including pregnancy test, urine toxicology, and electrocardiogram recordings.

Radiotracer

The standard (+)-McN 5652 was a gift from R.W. Johnson Pharmaceutical Research Institute. The precursor for the production of [11C] (+)-McN 5652, (+)-McN butyryl thioester tartrate, was prepared from (+)-McN 5652 by a modified literature procedure (28) as described previously (29). Radiochemical and chemical purity of [11C](+)–McN 5652 in saline was >95%.

PET Protocol

PET imaging was performed with the ECAT EXACT HR+ (Siemens/CTI, Knoxville, Tenn.). Sixty-three slices covered an axial field of view of 15.5 cm, axial sampling of 3.46 mm, three-dimensional mode in plane and axial resolution of 4.4 and 4.1 mm, respectively, full width at half maximum at the center of the field of view. An arterial catheter was inserted in the radial artery after completion of the Allen test and infiltration of the skin with 1% lidocaine. A venous catheter was inserted in a forearm vein on the opposite side. Head movement minimization was achieved with a polyurethane head immobilization system (Soule Medical, Tampa, Fla.) (30). A 10-minute transmission scan was obtained before radiotracer injection. [11C]McN 5652 was injected intravenously over 45 seconds. Emission data were collected in the three-dimensional mode for 120 minutes as 21 successive frames of increasing duration (three for 20 seconds, three for 1 minute, three for 2 minutes, two for 5 minutes, 10 for 10 minutes).

Input Function Measurement

Following radiotracer injection, arterial samples were collected every 10 seconds with an automated sampling system for the first 2 minutes, and manually thereafter at longer intervals. A total of 32 samples were obtained per scan. Seven samples (collected at 2, 16, 30, 50, 70, 90, and 120 minutes) were further processed by high-performance liquid chromatography to measure the fraction of plasma activity representing unmetabolized parent compound (18).

A biexponential function was fitted to the seven measured unmetabolized fractions, which was then used to interpolate values between the measurements. The smallest exponential of the unmetabolized fraction curve, λsat, was constrained to the difference between λer, the terminal rate of washout of cerebellar activity, and λ0,t, the smallest elimination rate constant of the total plasma activity (31).

The input function was calculated as the product of total counts and interpolated unmetabolized fraction at each time point. The measured input function values (Ca(t) [mCi/ml]) were fitted to a sum of three exponentials from the time of peak plasma activity, and the fitted values were used as the input to the kinetic analysis. The initial distribution volume (Vp [liters]) was calculated as the ratio of injected dose to peak plasma parent concentration. The clearance of the parent compound (CL [liters/hour]) was calculated as the ratio of the injected dose to the area under the curve of the input function (32).

The high retention (>90%) of free [11C]McN 5652 on the filter precludes the free fraction measurement of [11C]McN 5652; therefore, plasma f1 was not determined (18).

MRI Acquisition and Segmentation Procedures

MRIs, three-dimensional spoiled gradient-recall acquisition in the steady state, were acquired on a GE 1.5-T Signa Advantage system, as previously described (33). MRI segmentation was performed within Medx (Sensor Systems, Inc., Sterling, Va.), with original subroutines implemented in MATLAB (The Math Works, Inc., Natick, Mass.). Steps for MRI segmentation included correction for field inhomogeneities, fitting of the intensity distribution to a sum of three Gaussian functions, voxel classification, and post filtering (34).

Image Analysis

Images were reconstructed to a 128×128 matrix (pixel size of 2.5×2.5 mm²). Reconstruction was performed with attenuation correction that used the transmission data and a Shepp 0.5 filter (cutoff 0.5 cycles/projection ray). Reconstructed image files were then processed with the image analysis software Medx (Sensor Systems, Inc., Sterling, Va.). If indicated following visual inspection, frames were realigned to a frame of reference using a least-squares algorithm for within-modality coregistration (automated image registration) (35). The results of the frame-to-frame realignment were checked again visually. Following frame to frame registration, the 21 frames were summed to one dataset, which was coregistered to the MRI dataset using automated image registration (35). The spatial transformation derived from the summed PET registration procedure was then applied to each individual
frame. Thus, each PET frame was resampled in the coronal plane to a voxel volume of 1.5×0.9×0.9 mm³.

Regions of interest (N=10) and region of reference (cerebellum) boundaries were drawn on the MRI according to criteria derived from brain atlases (36, 37) and published reports (38–41). Analysis was restricted to regions of interest where serotonin transporter density is high enough to provide a reliable signal. Regions of interest included the midbrain (encompassing serotonin transporter dense structures such as the raphe nuclei, substantia nigra, locus ceruleus, ventral tegmental area, and superior and inferior colliculi), thalamus, dorsal caudate, dorsal putamen, ventral striatum, amygdala, entorhinal cortex, hippocampus, parahippocampal gyrus, and the anterior cingulate cortex. A segmentation-based method was used for the neocortical regions, and a direct identification method was used for the subcortical regions (33). For bilateral regions, right and left values were averaged. The contribution of plasma total activity to the regional activity was calculated assuming a 5% blood volume in the regions of interest (42), and tissue activities were calculated as the total regional activity minus the plasma contribution.

**Derivation of Distribution Volumes**

Derivation of [11C]McN 5652 regional tissue distribution volumes was performed with kinetic modeling using the arterial input function and a one-tissue compartment model. This model has been demonstrated to provide reliable estimates of total distribution volume for [11C]McN 5652 (16, 18, 19). Total distribution volume (VT [ml/g]), which is equal to the ratio of tissue to plasma parent activity at equilibrium, was derived as the K1/k2 ratio, where K1 (ml/g/min) and k2 (1/min) are the unidirectional fractional rate constants for the transfer of the tracer in and out of the brain, respectively (43, 44). Kinetic parameters were derived by nonlinear regression using a Levenberg-Marquardt least-squares minimization procedure (45) implemented in MATLAB (The Math Works, Inc., South Natick, Mass.) as previously described (43). Given the unequal sampling over time (increasing frame acquisition time from the beginning to the end of the study), the least-squares minimization procedure was weighted by the frame acquisition time.

**Serotonin Transporter Parameters**

Derivation of serotonin transporter parameters was based upon the following assumptions: 1) given the negligible density of serotonin transporter in the cerebellum (46–48), cerebellum total distribution volume was assumed to be representative of equilibrium nonspecific binding; 2) the nonspecific binding did not vary significantly between regions.

Two measures of serotonin transporter availability were calculated. The binding potential (ml/g) was derived as the difference in total distribution volume between the region of interest and the cerebellum, the reference region. The relationship between binding potential and serotonin transporter receptor parameters is given by (49)

\[ \text{BP} = \frac{f_1B_{max}}{K_D} \]

where \( B_{max} \) is the regional concentration of serotonin transporter (nmol/liter), and \( K_D \) is the in vivo affinity of the tracer for serotonin transporter (nmol/liter).

The specific-to-nonspecific equilibrium partition coefficient (\( V_3^{′′} \) [unitless]) was derived as the ratio of binding potential to total distribution volume in the cerebellum. The relationship between \( V_3^{′′} \) and serotonin transporter receptor parameters is given by (49)

\[ V_3^{′′} = \frac{f_2K_{max}}{K_D} \]

where \( f_2 \) is the free fraction of the nonspecific distribution volume in the brain (\( f_2 = f_1 \) / cerebellum VT).

**Statistical Analysis**

Between-group comparisons were assessed with unpaired two-tailed t tests with a probability of 0.05 set as the level of significance. The a priori hypothesis of this study related to serotonin transporter level in the anterior cingulate cortex, and therefore no correction for multiple comparisons was employed in the examination of this brain region. The other brain regions were also analyzed for between-group differences to determine the specificity of the difference seen in the anterior cingulate cortex. In order to correct for multiple comparisons, these analyses employed a univariate repeated-measures analysis of variance (ANOVA) with brain regions as the within-subject factor and diagnosis as the between-subject factor.

**Results**

**Group Composition**

Demographic data for study participants are shown in Table 1. No significant group differences were observed on any demographic factor. One individual in the impulsive

<table>
<thead>
<tr>
<th>Variable</th>
<th>Healthy Subjects (N=10)</th>
<th>Impulsive Aggression (N=10)</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>34</td>
<td>8</td>
<td>35</td>
<td>9</td>
</tr>
<tr>
<td>Injected dose (mCi)</td>
<td>11</td>
<td>6</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Injected mass (µg)</td>
<td>4.6</td>
<td>1.6</td>
<td>3.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Clearance (liters/hour)</td>
<td>143</td>
<td>31</td>
<td>159</td>
<td>55</td>
</tr>
<tr>
<td>Nonspecific distribution volume (ml/g)</td>
<td>19.3</td>
<td>3.8</td>
<td>20.3</td>
<td>2.0</td>
</tr>
</tbody>
</table>

As represented by total distribution volume in the cerebellum, the reference region due to negligible serotonin transporter density.
aggression group reported taking fluoxetine for 5 months several years before the study; all other participants were medication naive at the time of the scan. A significant degree of psychiatric comorbidity was present in the impulsive aggression group. This was particularly evident with regard to axis II diagnoses. All subjects had at least one personality disorder, the most common being borderline personality disorder (N=7), followed by paranoid (N=4), narcissistic (N=3), obsessive-compulsive (N=4), schizotypal (N=2), antisocial (N=2), histrionic (N=2), dependent (N=1), and avoidant (N=1) personality disorders. At the time of scanning, no subject had an active axis I mood disorder. However, several subjects were retrospectively diagnosed with comorbid mood illness. These included major depressive disorder (N=5), dysthymia (N=1), and bipolar II disorder (N=1). Anxiety disorders were also prevalent in this group, such as generalized anxiety disorder (N=1), adjustment disorder (N=1), social phobia (N=1), obsessive-compulsive disorder (N=1), and body dysmorphic disorder (N=1). In addition, there were two subjects with a history of alcohol dependence and two with a history of alcohol abuse; all had been in remission for greater than 6 months.

Scan Parameters

Scan parameters, including the injected dose, injected mass, and specific activity, did not differ between groups (Table 1). The plasma clearances of $[^{11}C]McN$ 5652 were similar between the groups. No significant group difference was observed in $[^{11}C]McN$ 5652 nonspecific distribution volume, measured as the cerebellar distribution volume.

Regional Volumes

Table 2 lists the average size for each region of interest in mm$^3$. There was no difference in anterior cingulate cortex volume between the groups ($t=0.18$, df=18, $p=0.86$). For the other regions, a repeated-measures ANOVA showed a significant effect of region ($F=121.4$, df=1, 21, $p<0.0001$), no effect of group ($F=0.04$, df=1, 18, $p=0.85$), and no group-by-region interaction ($F=0.23$, df=1, 21, $p=0.99$).

Region of Interest Kinetic Analysis

After injection, the accumulation of $[^{11}C]McN$ 5652 activity was consistent with the known distribution of serotonin transporter in the brain (Figure 1). The kinetic analysis converged in all regions for all subjects. Regional values of

---

FIGURE 1. PET and MRI Images of Serotonin Transporter Distribution in a 37-Year-Old Woman With Impulsive Aggression

---

$^{a}$ The PET images represent the activity from 40 to 90 minutes after injection of 13.76 mCi of $[^{11}C]McN$ 5652. Accumulation of activity can be seen in the thalamus and caudate in the midline sagittal plane images on the left. The region of interest for the anterior cingulate cortex is shaded in red on both the sagittal and transaxial MRI slices. The transaxial section demonstrates the high level of serotonin transporter in the striatum. The coronal plane images on the right show the ventral-dorsal serotonin transporter gradient in the striatum; the anterior cingulate cortex is not visible in the coronal section at this level.
the unidirectional fractional rate constant of tracer delivery as well as the regional total distribution volume values are shown in Table 2. There was no difference between groups in the tracer delivery values for the anterior cingulate cortex (t=−0.52, df=18, p=0.14). Examining all regions, a significant region effect was observed (repeated-measures ANOVA: F=153.0, df=1, 21, p<0.0001), with no difference between groups (repeated-measures ANOVA: F=0.06, df=1, 18, p=0.81) and no group-by-region interaction (repeated-measures ANOVA: F=0.06, df=1, 21, p=0.80). Similarly, for total distribution volume, no group differences were present in the anterior cingulate cortex (t=0.53, df=18, p=0.60). For the other regions, a region effect was observed (repeated-measures ANOVA: F=152.1, df=1, 21, p<0.0001), with no effect of group (repeated-measures ANOVA: F=0.09, df=1, 18, p=0.77) and no group-by-region interaction (repeated-measures ANOVA: F=0.60, df=1, 21, p=0.83).

**[11C]McN 5652 Binding Potential**

As shown in Table 3, a significant difference in binding potential between healthy subjects and subjects with impulsive aggression was revealed in the anterior cingulate cortex. For all other regions, no group differences were detected (repeated-measures ANOVA: F=1.24, df=1, 18, p=0.28). A significant regional effect was seen (F=137.1, df=1, 21, p<0.0001), with no group-by-region interaction (F=0.48, df=1, 21, p=0.90).

**[11C]McN 5652 Regional V_{3}'**

As shown in Table 4, a significant difference in V_{3}' between healthy subjects and subjects with impulsive aggression was revealed in the anterior cingulate cortex. Examination of all regions in a repeated-measure design resulted in no significant difference between the groups (F=2.83, df=1, 18, p=0.11), with a significant regional difference (F=163.6, df=1, 21, p<0.0001) and no region-by-group interaction (F=1.0, df=1, 21, p=0.43). The effect size of the reduction in [11C]McN 5652 V_{3}' in all regions examined is presented in Figure 2.

**Laterality**

Given the finding of reduced serotonin transporter parameters in the anterior cingulate cortex of patients relative to healthy comparison subjects, we performed a post
serotonin transporter in impulsive aggression

hoc analysis to determine if this finding was specific to the left or right hemisphere. Table 5 shows the results of this analysis. No difference was seen in total distribution volume in either the right or left anterior cingulate cortex in patients relative to healthy subjects. However, for the outcome measures binding potential and V3″, patients with impulsive aggression had significantly lower values in the left anterior cingulate cortex than did the healthy subjects, whereas no significant difference was seen in the right anterior cingulate cortex. This finding was not due to differences between the right and left anterior cingulate cortex in healthy subjects, since total distribution volume, binding potential, and V3″ values were not different in the left relative to the right anterior cingulate cortex. For patients with impulsive aggression, no differences in total distribution volume were seen when comparing the left anterior cingulate cortex with the right anterior cingulate cortex. However, for binding potential and V3″, the values in the left anterior cingulate cortex tended to be lower than the right anterior cingulate cortex (paired t test, p=0.06). No significant differences were observed for any other region in this analysis.

Discussion

This study suggests that pathological impulsive aggression is associated with a reduction in serotonin transporter availability in the anterior cingulate cortex, a reduction that might reflect reduced 5-HT innervation. Although not statistically significant, our data also suggest a modest decrease in serotonin transporter availability in other regions.

The use of a fully quantitative imaging method was a strength of this study. The kinetic analysis is not affected by potential group differences in radiotracer plasma clearance or regional cerebral blood flow (50). Measurement of the arterial input function enabled the quantitative derivation of distribution volumes, from which both binding potential and V3″ were calculated. Although methods for deriving [11C]McN 5652 V3″ without arterial sampling have been proposed (20), only with the arterial input function could we demonstrate the absence of group differences in cerebellum distribution volumes and thus validate the use of V3″ for between-group comparisons of receptor parameters. In the absence of group differences in cerebellum distribution volumes, results derived with binding potential and V3″ (i.e., binding potential normalized by the nonspecific distribution volume) are essentially similar, which was the case here.

This study has several limitations. The small size of this study group was adequate to detect the relatively large decrease in serotonin transporter in the anterior cingulate cortex in subjects with impulsive aggression but limited our power to detect potential significant differences in other regions. As described, in all regions, V3″ values were
lower for the subjects with impulsive aggression than the healthy subjects. Analysis of a larger group would be required to further explore this difference. Examination of the overall effect size for the repeated-measures ANOVA reveals that approximately 30 individuals per group would be required to detect this difference at the level of \( p < 0.05 \). The effect size (\( d \)) for the differences in the individual regions ranged from 1.22 in the anterior cingulate cortex to 0.16 in the entorhinal cortex.

Another limitation of this study arises from the use of \( [11C] \)McN 5652 to measure serotonin transporter. This radiotracer is associated with high levels of nonspecific binding, limiting the ability to quantify serotonin transporter in regions of low serotonin transporter density such as the neocortex (18). The ability to detect differences in these regions will likely be enhanced by the use of newly developed radiotracers for serotonin transporter such as \( [11C]3 \)-amino-4-[2-\{\[dimethylamino\]methyl\}[phenylthio]benzonitrile (\( [11C] \)DASB) (51) and \( [11C]2\{-[dimethylaminomethyl][phenylthio]\}5\)-fluoromethylphenylamine (\( [11C] \)AFM) (52, 53). Both tracers markedly improve the signal-to-noise ratio compared with \( [11C] \)McN 5652 (29, 53).

Given previous studies implicating the anterior cingulate cortex in impulsive aggression, this area was the primary focus of our study. The finding of a reduction in serotonin transporter density in this region is consistent with results from studies of cerebral metabolism in impulsive aggression. In response to serotonergic challenges, including \( \text{d,1-fenfluramine} \) and \( \text{m-CPP} \), the relative glucose metabolic rate of individuals with impulsive aggression is blunted in the anterior cingulate cortex relative to healthy subjects (13, 14). Fenfluramine acts by causing a direct release of serotonin and antagonizing its reuptake, whereas \( \text{m-CPP} \) acts as a partial agonist at the postsynaptic 5-HT(2A) and 5-HT(2C) receptors. On the basis of these mechanisms of action alone it is not possible to comment on the locus of the abnormality in impulsive aggression, i.e., whether the blunted response is secondary to a pre- or postsynaptic problem. Our study provides some insight into this question, suggesting that a presynaptic deficiency in serotonin innervation exists in the anterior cingulate cortex in this disorder. This result is in line with the observations from studies of serotonin metabolites that have linked reduced serotonin markers with impulsive aggression (1, 3, 54). Our findings are also in agreement with the postmortem finding of a decrease in serotonin transporter binding in suicide victims (55), viewed as a specific form of self-directed impulsive/aggressive behavior, as well as the finding of reduced platelet serotonin transporter in aggression (56).

Davidson et al. (57) proposed that disruption of the normal regulation of emotion—which involves several interconnected brain regions including the orbitofrontal cortex, amygdala, and anterior cingulate cortex—plays a role in the generation of violence. The anterior cingulate cortex can be separated into a dorsal “cognitive” portion and a rostral-ventral “affective” region (58). Evidence from a variety of domains indicates that the affective subdivision of the anterior cingulate cortex regulates the intensity of response to emotional stimuli. For example, stimulation of this area in animal models increases the latency of attack behavior (59). In humans, PET studies of cerebral blood flow demonstrate activation of the ventral anterior cingulate cortex when anger is induced in healthy men using imagery (60). This same area is activated when symptoms are provoked in individuals with simple phobia, OCD, or PTSD (61–63). Further evidence that the rostral-ventral anterior cingulate cortex plays a role in emotional regulation comes from the finding that this area was activated when men attempted to suppress sexual arousal in response to erotic film excerpts but not in the nonsuppression condition (64). Given the aforementioned limitations of the radiotracer, we were unable to separate the anterior cingulate cortex into the dorsal and rostral-ventral components. However, we did examine the laterality of our finding. Previously, there have been no reports of aggression-related laterality in the anterior cingulate cortex, although work by our group has demonstrated blunted metabolic response to \( \text{m-CPP} \) in the left orbitofrontal cortex (14). The results of this study, indicating a left-sided predominance of the abnormality, are consistent with findings showing that traumatic brain lesions to the left frontal cortex give rise to aggression and hostility, whereas right-sided lesions lead to anxiety/depression (8).

Our finding of an abnormality of the serotonergic innervation in the anterior cingulate cortex in subjects with impulsive aggression is consistent with the hypothesis that alterations in the normal function of this area may lead to difficulties with affect and impulse modulation, resulting in increased impulsivity and aggression.

**Conclusions**

This study detected a significant difference in anterior cingulate cortex serotonin transporter availability in individuals with impulsive aggression. These individuals react aggressively in response to emotionally laden interpersonal stimuli, such as conflict or perceived disrespect. This is hypothesized to result from impaired regulation of negative emotions, believed to be one of the primary functions of the anterior cingulate cortex (57). Future work extending the results of this study to other brain regions involved in the regulation of emotion and behavioral responses, such as the orbitofrontal cortex, is warranted to further characterize alterations of 5-HT function in impulsive aggressivity.
References

32. Abi-Dargham A, Laurelle M, Selby J, Rattner Z, Baldwin RM, Zoghbi SS, Zoa-Ponce Y, Bremner JD, Hyde TM, Charney DS,


Objective: In an ongoing molecular genetic study of temperament, participants were genotyped to examine the association of smoking with two polymorphisms of the serotonin transporter gene (SERT): the promoter region, 5-HTTLPR, and an intronic variable-number-of-tandem-repeats region (VNTR).

Method: Full information was available for 330 families, and 244 “ever smokers” were identified (54 past smokers, 190 current smokers). The average number of cigarettes smoked per day was 13.12, and the mean Fagerstrom Tolerance Questionnaire score was 4.79. Associations of genotype, Tridimensional Personality Questionnaire scores, and smoking phenotype were tested by using a robust family design with a variance-components framework and by case-control analysis.

Results: There was a significant excess of the 5-HTTLPR long allele with the 12-repeat VNTR in current smokers, past smokers, and ever smokers, compared to participants who had never smoked. The results from the population design were confirmed in the family-based analysis. No association was observed between two quantitative measures of smoking and the polymorphisms. A weak association was observed between novelty seeking and the VNTR polymorphism and between reward and 5-HTTLPR. Smokers, regardless of gender, scored significantly higher on novelty seeking and did not differ on harm avoidance or reward.

Conclusions: There was a highly significant association between SERT and the categorical definition of smoking, irrespective of dependence level, suggesting that this gene influences the initiation of smoking. Mediation analysis failed to substantiate the hypothesis that novelty seeking partially mediates the effect of SERT on smoking. SERT appears to independently contribute to novelty seeking and smoking.

Genetic factors are important in determining the complex smoking phenotype, as shown by twin studies showing that genes partially confer susceptibility to nicotine dependence (1). As well as two genome scans (2, 3), there have been a number of tests of candidate genes (4), including the promoter region of the serotonin transporter gene (SERT) (4–9). The SERT promoter region is a 44-base-pair insertion/deletion polymorphism that was originally shown to be associated with anxiety-related personality traits (10). Indeed, an interaction of the short promoter polymorphism, anxiety-related personality traits, and smoking was observed in two studies (7, 8). Conversely, in a Japanese study it was the long promoter variant that was observed to be associated with smoking (6, 9).

We have now recruited a new group of 330 families, including 244 past and present smokers, in the framework of our continuing studies of normal personality (11), positioning us to validate previous studies showing an interaction between SERT, personality traits, and smoking behavior (7, 8). Toward this end, we genotyped this group for two SERT polymorphisms, the serotonin (5-HT) transporter linked promoter region (5-HTTLPR) and an intronic variable-number-of-tandem-repeats region (VNTR) (12). Both common repeats of the VNTR are purported to increase transcription of this gene (13). The association between smoking and SERT was analyzed both by using a case-control design and by a more robust family-based approach. Scoring allelic transmission in families (14, 15) allowed us to test association in this heterogeneous group by using both categorical and quantitative definitions of smoking.

Method

Subjects

Participants in our ongoing studies of personality (11), who are primarily but not exclusively college students at various locations in Israel and their families, are recruited by word of mouth and advertisements on campuses and at other institutions. The study group analyzed comprised 330 families, each consisting of two biological parents and two or more same-sex siblings. The subgroup of smokers in this nonclinical group was defined as participants who had smoked for at least 1 year, irrespective of level of dependence or number of cigarettes smoked. They were identified by a set of questions that subjects in the study answered, e.g., “Do you currently smoke?” and “Did you smoke in the past?”

There were 244 participants between the ages of 14 and 70 years (mean=28.91) who had ever smoked (“ever smokers”); 54 had
TABLE 1. Relation of Smoking Phenotype to Allele Frequency of Two Polymorphisms of the Serotonin Transporter Gene (SERT) in Nonclinical Subjects

<table>
<thead>
<tr>
<th>SERT Polymorphismb</th>
<th>Subjects Who Never Smoked (N=486)</th>
<th>Current Smokers (N=190)</th>
<th>Past Smokers (N=54)</th>
<th>Total Number of Alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Alleles</td>
<td>%</td>
<td>Number of Alleles</td>
<td>%</td>
</tr>
<tr>
<td>10-repeat VNTR</td>
<td>272</td>
<td>100.00</td>
<td>90</td>
<td>100.00</td>
</tr>
<tr>
<td>5-HTTLPR long allele</td>
<td>224</td>
<td>82.35</td>
<td>78</td>
<td>86.67</td>
</tr>
<tr>
<td>5-HTTLPR short allele</td>
<td>48</td>
<td>17.65</td>
<td>12</td>
<td>13.33</td>
</tr>
<tr>
<td>12-repeat VNTR</td>
<td>700</td>
<td>100.00</td>
<td>290</td>
<td>100.00</td>
</tr>
<tr>
<td>5-HTTLPR long allele</td>
<td>231</td>
<td>33.00</td>
<td>132</td>
<td>45.52</td>
</tr>
<tr>
<td>5-HTTLPR short allele</td>
<td>469</td>
<td>67.00</td>
<td>158</td>
<td>54.48</td>
</tr>
</tbody>
</table>

a Subjects must have smoked for at least 1 year to be considered smokers.
b VNTR, intronic variable-number-of-tandem-repeats region.
c QTDT (quantitative transmission disequilibrium tests) is used to analyze quantitative or discrete traits, either in nuclear families, with or without parental genotypes, or in extended pedigrees to analyze quantitative or discrete traits. Models for association, linkage, and linkage and association were employed.

TABLE 2. Family-Based Test of Association Between Smoking Phenotype and Two Polymorphisms of the Serotonin Transporter Gene (SERT) in Nonclinical Subjects

<table>
<thead>
<tr>
<th>Group and SERT Polymorphismb</th>
<th>Results of Family-Based Test of Linkage Disequilibrium (QTDT)c</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Null Model</td>
<td>Full Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>df</td>
<td>Log Likelihood</td>
<td>df</td>
<td>Log Likelihood</td>
<td>χ² (df=1)</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>Subjects Who Never Smoked (N=486)</td>
<td>594</td>
<td>282.40</td>
<td>593</td>
<td>276.01</td>
<td>12.79</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>Current smokers (N=190)</td>
<td>594</td>
<td>282.60</td>
<td>593</td>
<td>276.44</td>
<td>12.33</td>
<td>0.0004</td>
<td></td>
</tr>
<tr>
<td>10-repeat VNTR</td>
<td>594</td>
<td>352.32</td>
<td>593</td>
<td>342.59</td>
<td>19.45</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>12-repeat VNTR</td>
<td>594</td>
<td>352.33</td>
<td>593</td>
<td>342.70</td>
<td>19.26</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>Past smokers (N=54)</td>
<td>594</td>
<td>112.96</td>
<td>593</td>
<td>110.96</td>
<td>4.00</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>10-repeat VNTR</td>
<td>594</td>
<td>113.03</td>
<td>593</td>
<td>110.93</td>
<td>4.19</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

a Subjects must have smoked for at least 1 year to be considered smokers.
b VNTR, intronic variable-number-of-tandem-repeats region.
c Two models were evaluated with QTDT (quantitative transmission disequilibrium tests) (14, 15). In the full model, means = mu + COVARIATE_HTTLPR + AGE + BMI + Tridimensional Personality Questionnaire novelty seeking score + B, with variances = Ve + Vg + Va. In the null model, means = mu + COVARIATE_HTTLPR + AGE + BMI + Tridimensional Personality Questionnaire novelty seeking score + B + W, with variances = Ve + Vg + Va. Httlpr, allele for the serotonin transporter linked promoter region; BMI, body mass index; B, between-family component of association; W, within-family component of association; e, nonshared environment; g, polygenic effects (function of relatedness between family members, perhaps due to polygenes); a, additive genetic effects. The chi-square value for determining the overall global p value was derived by subtracting the chi-square value of the full model from the chi-square value of the null model. Alleles with rare frequencies of 0.05% or less were lumped together. The individual effects for all other alleles were estimated, producing a single, global p value. A fuller explanation of these models can be found at http://www.sph.umich.edu/csg/abecasis/ and in references 14 and 15.

Genotyping

DNA was obtained from all family members, including parents. DNA was extracted with the Master Pure kit (Epicentre, Madison, Wis.). The 5-HTTLPR polymorphism was characterized as previously described by our laboratory (18).

Statistical Procedures

QTDT (quantitative transmission disequilibrium tests) is used to analyze quantitative or discrete traits, either in nuclear families, with or without parental genotypes, or in extended pedigrees (14, 15). Since simple models of association do not provide valid tests of linkage disequilibrium when multiple offspring per family are considered, we used the robust variance-components procedures as detailed in the QTDT software package. Models for association, linkage, and linkage and association were employed.

Results

In the first analysis we used a population-based case-control design to examine the association between SERT polymorphisms (5-HTTLPR and intronic VNTR) and the smoking phenotype (Table 1). There was a significant excess of the 5-HTTLPR long allele with the 12-repeat VNTR (Pearson χ²=15.57, df=2, p=0.0004) in both current and past smokers, compared to people who had never smoked smoked in the past, and 190 were current smokers. For current smokers, the average number of cigarettes smoked per day was 13.12 (SD=8.65, range=1–50), and their mean score on the Fagerstrom Tolerance Questionnaire (16) was 4.79 (SD=2.37). The body mass index of the current smokers was 22.16 (SD=2.18). Only 25% of the smokers had a Fagerstrom Tolerance Questionnaire score of 6 or above, one common measurement of nicotine dependence. Similarly, only 10% of the smokers in this study smoked more than 20 cigarettes per day, and “heavy smoking” is commonly defined as more than 30 or more than 40 cigarettes per day. The lifetime duration of smoking was 6.4 years (range=1–17 years, with a minimum of 1 year of smoking). However, only 8% of the smokers had smoked for more than 10 years, reflecting the young age of this group. Overall, the smokers in this study were not “heavy smokers,” and only a minority were clearly “dependent,” i.e., had a Fagerstrom Tolerance Questionnaire score of 6 or higher.

Each contact person received a number of booklets (equal to the number of participating siblings), containing the Tridimensional Personality Questionnaire (17) and other self-report questionnaires, and two sterile test tubes per family member, each containing 10 cm³ of Aquafresh mouthwash, for DNA sampling. The booklets, completed by the siblings, and the DNA mouthwashes were returned by mail or hand delivered to an office. Only siblings completed the questionnaires. The contact person received a modest monetary incentive for the family’s participation.

The study was approved by the local institutional review board and by the Israeli Ministry of Health Genetics Committee.

KREMER, BACHNER-MELMAN, RESHEF, ET AL.

Am J Psychiatry 162:5, May 2005
http://ajp.psychiatryonline.org

925
Since population-based case-control designs are prone to errors due to population stratification, we also examined these findings using a robust family-based QTDT analysis. The results from the population design were confirmed in the family-based analysis (Table 2) for current smokers, ever smokers, and past smokers. Again, no association was observed between current smokers, ever smokers, and past smokers.’é current smokers and past smokers scored significantly higher on the Tridimensional Personality Questionnaire trait of novelty seeking than did never smokers. There were no significant differences between past and current smokers. Smokers did not differ from never smokers on either the reward or harm avoidance measure. Smokers scored nonsignificantly lower on persistence. As observed previously in the Israeli population (20), the women scored higher than the men on both harm avoidance and reward, but we found no significant interaction between sex and smoking phenotype regarding the Tridimensional Personality Questionnaire temperament factors. Both male and female smokers scored higher on novelty seeking than sex-matched never smokers according to analysis of variance.

We next examined the relationships among the SERT polymorphisms, personality traits, and smoking phenotype. Although a strong association was observed between the SERT variants and the smoking phenotype (Table 1), only weak relationships were observed between the polymorphisms and personality traits according to the QTDT family-based analysis (Table 4). No association was detected in this study group between the SERT polymorphisms and harm avoidance, whereas a weak but significant association was observed between novelty seeking and the VNTR polymorphism. Additionally, a weak associ-

### TABLE 3. Personality Traits of Smokers and Nonsmokers Among Male and Female Nonclinical Subjectsa

<table>
<thead>
<tr>
<th>Personality Trait and Smoking Phenotype</th>
<th>Score on Tridimensional Personality Questionnaire</th>
<th>Difference Between Sexes</th>
<th>Difference Between Ever Smokers (current and past) and Never Smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
<td>F (df=1, 776)</td>
</tr>
<tr>
<td>Novelty seeking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smokers</td>
<td>180</td>
<td>337</td>
<td>0.63</td>
</tr>
<tr>
<td>Current smokers</td>
<td>84</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Past smokers</td>
<td>22</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Harm avoidance</td>
<td></td>
<td></td>
<td>6.56</td>
</tr>
<tr>
<td>Never smokers</td>
<td>180</td>
<td>337</td>
<td></td>
</tr>
<tr>
<td>Current smokers</td>
<td>84</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Past smokers</td>
<td>22</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Reward</td>
<td></td>
<td></td>
<td>35.00</td>
</tr>
<tr>
<td>Never smokers</td>
<td>180</td>
<td>337</td>
<td></td>
</tr>
<tr>
<td>Current smokers</td>
<td>84</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Past smokers</td>
<td>22</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Never smokers</td>
<td>180</td>
<td>337</td>
<td></td>
</tr>
<tr>
<td>Current smokers</td>
<td>84</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Past smokers</td>
<td>22</td>
<td>35</td>
<td></td>
</tr>
</tbody>
</table>

a Subjects must have smoked for at least 1 year to be considered smokers. There was no significant sex-by-smoking interaction for any of the personality traits.

b Significant (p<0.05) after Bonferroni correction.

("never smokers"). When ever smokers (past and current) were compared to never smokers (data not shown), there was also a significant excess of the 5-HTTLPR long allele with the 12-repeat VNTR (χ²=14.38, df=1, p=0.0001), and the estimated risk of these polymorphisms occurring together in ever smokers was 1.37 (95% confidence interval, 1.17–1.61).

We also stratified the smokers by scores on the Fagerstrom Tolerance Questionnaire and by the number of cigarettes smoked per day. Within the group of current smokers, no association was observed between Fagerstrom Tolerance Questionnaire score (1–5 versus >5) and SERT (10-repeat VNTR: χ²=0.61, df=1, n.s.); 12-repeat VNTR: χ²=0.05, df=1, n.s.; with either the short or long 5-HTTLPR allele, respectively). Additionally, no association was observed between the number of cigarettes smoked per day (0–5, 6–15, >15) and SERT (10-repeat VNTR: χ²=0.58, df=2, n.s.; 12-repeat VNTR: χ²=2.40, df=2, n.s.; with either the short or long 5-HTTLPR allele, respectively).

Since population-based case-control designs are prone to errors due to population stratification, we also examined these findings using a robust family-based QTDT analysis. The results from the population design were confirmed in the family-based analysis (Table 2) for current smokers, ever smokers, and past smokers. Again, no association was observed between current smokers, ever smokers, and past smokers. As observed previously in the Israeli population (20), the women scored higher than the men on both harm avoidance and reward, but we found no significant interaction between sex and smoking phenotype regarding the Tridimensional Personality Questionnaire temperament factors. Both male and female smokers scored higher on novelty seeking than sex-matched never smokers according to analysis of variance.
TABLE 4. Family-Based Test of Associations Among Smoking Phenotype, Personality Traits, and Two Polymorphisms of the Serotonin Transporter Gene (SERT) in 244 Nonclinical Subjects

<table>
<thead>
<tr>
<th>Tridimensional Personality Questionnaire Trait and SERT Polymorphism&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Results of Family-Based Test of Linkage Disequilibrium (QTDT)&lt;sup&gt;b&lt;/sup&gt;</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Null Model</td>
<td></td>
<td>Full Model</td>
<td></td>
<td>Chi-Square Analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>df</td>
<td>Log Likelihood</td>
<td>df</td>
<td>Log Likelihood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harm avoidance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-HTTLPR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long allele</td>
<td>734</td>
<td>2385.77</td>
<td>733</td>
<td>2385.01</td>
<td>1.50</td>
<td>0.22</td>
</tr>
<tr>
<td>Short allele</td>
<td>734</td>
<td>2385.77</td>
<td>733</td>
<td>2385.31</td>
<td>0.91</td>
<td>0.34</td>
</tr>
<tr>
<td>VNTR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-repeat</td>
<td>727</td>
<td>2360.58</td>
<td>726</td>
<td>2360.52</td>
<td>0.12</td>
<td>0.73</td>
</tr>
<tr>
<td>12-repeat</td>
<td>727</td>
<td>2360.52</td>
<td>726</td>
<td>2360.45</td>
<td>0.14</td>
<td>0.70</td>
</tr>
<tr>
<td>Novelty seeking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-HTTLPR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long allele</td>
<td>608</td>
<td>1895.23</td>
<td>607</td>
<td>1894.69</td>
<td>1.09</td>
<td>0.29</td>
</tr>
<tr>
<td>Short allele</td>
<td>608</td>
<td>1895.15</td>
<td>607</td>
<td>1894.71</td>
<td>0.88</td>
<td>0.35</td>
</tr>
<tr>
<td>VNTR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-repeat</td>
<td>601</td>
<td>1869.27</td>
<td>600</td>
<td>1867.28</td>
<td>3.97</td>
<td>0.05</td>
</tr>
<tr>
<td>12-repeat</td>
<td>601</td>
<td>1869.26</td>
<td>600</td>
<td>1867.04</td>
<td>4.45</td>
<td>0.04</td>
</tr>
<tr>
<td>Reward</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-HTTLPR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long allele</td>
<td>608</td>
<td>1689.56</td>
<td>607</td>
<td>1687.26</td>
<td>4.59</td>
<td>0.04</td>
</tr>
<tr>
<td>Short allele</td>
<td>608</td>
<td>1689.56</td>
<td>607</td>
<td>1687.34</td>
<td>4.44</td>
<td>0.04</td>
</tr>
<tr>
<td>VNTR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-repeat</td>
<td>601</td>
<td>1671.46</td>
<td>600</td>
<td>1671.45</td>
<td>0.01</td>
<td>0.92</td>
</tr>
<tr>
<td>12-repeat</td>
<td>601</td>
<td>1671.62</td>
<td>600</td>
<td>1671.62</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<sup>a</sup>Subjects must have smoked for at least 1 year to be considered smokers. Both current and past smokers were classified as “ever smokers.”<br><sup>b</sup>VNTR, intronic variable-number-of-tandem-repeats region. 5-HTTLPR, serotonin transporter linked promoter region.<br><sup>c</sup>Two models were evaluated with QTDT (quantitative transmission disequilibrium tests) (14, 15). In the null model, means = μ + COVARIATE_AGE + BMI + EVER SMOKER + B, with variances = Ve + Vg + Va. In the full model, means = μ + COVARIATE_AGE + BMI + EVER SMOKER + B + W, with variances = Ve + Vg + Va. BMI, body mass index; B, between-family component of association; W, within-family component of association; e, nonshared environment; g, polygenic effects (function of relatedness between family members, perhaps due to polygenes); a, additive genetic effects. The chi-square value for determining the overall global p value was derived by subtracting the chi-square value of the null model from the chi-square value of the full model. Alleles with rare frequencies of 0.05% or less were lumped together. The individual effects for all other alleles were estimated, producing a single, global p value. A fuller explanation of these models can be found at http://www.sph.umich.edu/csg/abecasis/ and in references 14 and 15.

The main finding of the current study was a highly significant association between SERT and smoking. The long 5-HTTLPR allele with the 12-repeat VNTR conferred modest risk for smoking according to both case-control analysis and the more robust family-based design. Two previous Japanese studies (6, 9) also showed an association between the long 5-HTTLPR allele and smoking by means of a case-control design.

Many North American studies have shown that tobacco use is associated with several related personality features among adolescents and adults, including extraversion, impulsivity, risk taking, sensation seeking, monotony avoidance, novelty seeking and rebelliousness, and psychopathic and antisocial personality (23). Additionally, longitudinal studies have revealed that many of the aforementioned personality characteristics predict tobacco use (24). However, anxiety-related personality traits are also predictors of smoking behavior (25). For example, panic attacks are associated with greater risk of cigarette smoking, and neuroticism may play an essential role in this relationship (26). It appears that no single personality type generates risk for this complex phenotype, and either extraversion and neuroticism, depending on the cultural and genetic milieu, may contribute to initiation and persistence of smoking. The most notable difference between the current investigation and the American studies (5, 7, 8) is ethnicity and cultural setting, although, it should be noted, three different personality inventories were also employed, the Tridimensional Personality Questionnaire (17), Eysenck Personality Inventory (27), and NEO Personality Inventory, Revised (28). We believe it is unlikely that a difference due to the specific personality questionnaire...
used explains the contrasting results, since the personality traits of novelty seeking and harm avoidance and the NEO equivalents (extraversion and neuroticism) correlate well in most studies (29). It is not surprising that the particulars of the interaction between heritable personality traits, a complex behavioral phenotype such as smoking, and SERT differ across cultural and ethnic categories. It is intriguing that in two North American studies (5, 7, 8), in which smokers scored high on neuroticism, the short promoter variant showed an association with this phenotype, whereas in an Israeli population, in which smokers scored high on extraversion or sensation seeking, the long promoter variant was associated with smoking. Thus, the apparently opposing findings regarding which SERT promoter region allele is associated with smoking in two diverse cultural and ethnic groups are resolved by considering which personality trait (novelty/sensation seeking versus neuroticism) characterizes smokers. The role of the short allele in the North American study and the involvement of the long allele in the Israeli study therefore make biological and psychological sense.

The degree of nicotine dependence is unlikely to account for the particular SERT variant associated with smoking, since smoking habits vary considerably across studies without regard to genotype (5–9). The study by Ler- man et al. (5, 7), which showed an association with the short SERT variant, included smokers who had smoked at least five cigarettes per day for at least 1 year and were likely interested in quitting. Hu et al. (8) recruited subjects who had smoked an average of 20 cigarettes per day for an average of 15 years, and in this group an association with the short SERT genotype was observed. The Japanese sub- jects consumed more than 25 cigarettes daily, and an asso- ciation with the long SERT variant was observed. In the current investigation, the subjects smoked an average of 13 cigarettes daily and there were few “heavy” smokers, and again, an association with the long variant was observed.

We used a broad definition of smoking and studied a range of smoking phenotypes, raising the question of whether subjects within this heterogeneous group can be meaningfully compared. In the genetic analysis, however, the smoking phenotype was examined not only as a categorical trait but also as a quantitative trait. The powerful quantitative approach allowed us to test in the genetic model the degree of smoking dependence measured by number of cigarettes per day and score on the Fagerstrom Tolerance Questionnaire. However, neither quantitative measure showed any association with SERT. The highly sig- nificant association observed between the categorical defi- nition of smoking and SERT, irrespective of dependence, suggests that this gene is more involved in the initiation of smoking than in its persistence or the level of dependence.

Short-term administration of nicotine releases seroto- nin, whereas long-term treatment depletes brain seroto- nin, and there is strong evidence that serotonergic tone plays a permissive role in the expression of nicotine’s ef- fects (30). Subjects with low serotonergic tone due to the presence of the long promoter variant may be particularly sensitive to the serotonin-releasing effects of short-term nicotine administration and thus at greater risk of develop- ing dependence on cigarettes. Once initiated, smoking will further reduce serotonin levels, thus exacerbating the dependence on nicotine in subjects with a more transcriptionally efficient transporter gene.

Similar to other addictive drugs, nicotine is thought to affect the brain’s dopaminergic reward system, which in- cludes parts of the nucleus accumbens and amygdala.
short-term reinforcing effects of drugs of abuse involve associated with dopaminergic activity, key elements for the brain reward system in rats (35). Although the brain reward system is often associated with dopaminergic functioning. Indeed, serotonergic activity has the 5-HT2C receptor in mediating mesolimbic dopamine functioning. Indeed, serotonergic activity has a dual effect on stimulation of the brain reward system in rats (35). Although the brain reward system is often associated with dopaminergic activity, key elements for the short-term reinforcing effects of drugs of abuse involve other neurotransmitters, such as opioid peptides, γ-aminobutyric acid, glutamate, and serotonin (36).

In the present study, only a weak association was observed between novelty seeking and SERT, principally with the intronic VNTR, and despite the evidence that impulsive behavior is a risk factor for smoking in the Israeli population we studied, mediation analysis (20, 21) carried out on the current data does not support the hypothesis that novelty seeking mediates the effect of SERT on smoking. SERT independently contributes weakly to novelty seeking and more strongly to the smoking phenotype. Further studies across cultural and ethnic groups are required to clarify the behavioral pathways that mediate the effect of SERT on smoking.

References


35. Harrison AA, Markou A: Serotonergic manipulations both potentiate and reduce brain stimulation reward in rats: involvement of serotonin-1A receptors. J Pharmacol Exp Ther 2001; 297:316–325
Thalamic and Prefrontal FDG Uptake in Never Medicated Patients With Schizophrenia

Douglas S. Lehrer, M.D.
Bradley T. Christian, Ph.D.
Joseph Mantil, M.D., Ph.D.
Aaron C. Murray, B.A.
Bradley R. Buchsbaum, Ph.D.
Terrence R. Oakes, Ph.D.
William Byne, M.D., Ph.D.
Eileen M. Kemether, M.D.
Monte S. Buchsbaum, M.D.

Objective: Because neuroleptic treatment may cause long-lasting changes in brain structure and function, a group of patients with schizophrenia who had never been medicated was recruited to examine regional glucose metabolic rates in the frontal-striato-thalamic circuit.

Method: Twelve never medicated patients with schizophrenia (seven men, five women; mean age=29 years) and 13 normal volunteers (eight men and five women; mean age=28.5 years) underwent 18F-fluorodeoxyglucose (FDG) positron emission tomography, and coregistered anatomical magnetic resonance imaging scans were also obtained. During FDG uptake, subjects performed a spatial attention task previously shown to activate the pulvinar region of the thalamus.

Results: Diminished regional glucose metabolism was found in the medial dorsal nucleus, posterior thalamus, and prefrontal cortex of patients with schizophrenia relative to normal volunteers, extending earlier results from studies of medicated and previously medicated patients.

Conclusions: The finding of lower relative metabolic rates in the frontothalamic circuits of patients with schizophrenia is consistent with extended circuit deficits involving interactions of frontal executive areas with thalamic sensory and association processes.

The prefrontal cortex, striatum, and thalamus form a neural circuit important in regulating sensory input, attention, and action. Deficits in these three areas in schizophrenia have been widely reported in both structural (1) and functional (2) brain imaging studies. The thalamus comprises multiple nuclei that relay and filter sensory and higher-order inputs to and from the cerebral cortex and limbic structures. Thus, it is a candidate structure for abnormality in schizophrenia, a disease that includes disturbed sensory and attentional function (3). Especially important within the thalamus may be the medial dorsal nucleus (with its interconnections with the prefrontal cortex) and the pulvinar (also interconnected with frontal and temporal regions). Both nuclei interact with the cortex in high-level cognitive activity and have been found to have reduced volumes in schizophrenia in postmortem as well as magnetic resonance imaging (MRI) studies (4–6).

In our earliest positron emission tomography (PET) studies, which used 18F-fluorodeoxyglucose (FDG) and employed older-generation PET scanners with lesser in-plane resolutions of 15 mm (7) or 7.5 mm (8), we did not find decreased metabolic rates in the thalamus of unmedicated (previously medicated) schizophrenia patients compared with healthy subjects. In a study of never medicated schizophrenia subjects, we found less prominent regional glucose metabolic rate (rGMR) in the mediodorsal nucleus region (6). In another study of never medicated patients (9), blood flow decreases in the thalamus, frontal lobe, and temporal lobe were identified. With increased spatial PET resolution and coregistered structural MRI (1.2-mm thick slices) scans obtained for the entire thalamic volume, lower rGMR in the medial dorsal nucleus of unmedicated schizophrenia patients relative to healthy subjects was statistically confirmed, although whole thalamus metabolism did not differ between groups (10). However, when the nuclei were traced on coregistered MRI templates, decreased metabolic rates in the medial dorsal and centromedian nuclei (11), but not the pulvinar, were confirmed. Decreased rGMR in the remainder of the thalamus was again not found, indicating specificity of the effect to the medial association regions. That study, while large (61 normal subjects and 40 patients), was limited in having patients who, although unmedicated, had previously been treated with neuroleptics, which might have affected thalamic volume and rGMR. In addition, the uptake condition was the serial verbal learning task, which activates the frontal lobe but is not known to activate the pulvinar.

In the current study we selected a specific task known to activate the pulvinar (12) for the uptake condition and recruited patients who had never been medicated. We have undertaken two separate, independent coregistration methods to replicate thalamic effects and used internal brain fiducial landmarks to assess coregistration accuracy and alignment in the 12-point affine transformation to standard space.
Method

Subjects

A group of 12 psychotic patients (seven men and five women; mean age=29.0 years [SD=9.8, range=20–46]) was recruited from the greater Dayton, Ohio, area and was evenly divided between inpatients and outpatients. After complete description of the study, all subjects completed a verbal “informed consent post test.” All participants gave written informed consent. Subjects were evaluated with the Comprehensive Assessment of Symptoms and History (13), 18-item version of the Brief Psychiatric Rating Scale (BPRS) (14), and the Abnormal Involuntary Movement Scale (AIMS) (15) and were given their diagnosis by a psychiatrist (D.S.L.) who used DSM-IV criteria. Patients had either never been medicated (N=11) or almost never medicated (one subject had a lifetime neuroleptic exposure of no more than five doses). Patients had either never been medicated (N=11) or almost never medicated (one subject had a lifetime neuroleptic exposure of no more than five doses)

Table 1. Clinical Characteristics of a Group of Never Medicated Patients With Schizophrenia Recruited for a PET Study of Regional Glucose Metabolism in the Fronto-Striato-Thalamic Circuit

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Schizophrenia Patients (N=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary DSM-IV diagnosis</td>
<td>N</td>
</tr>
<tr>
<td>Schizophrenia, paranoid type</td>
<td>4</td>
</tr>
<tr>
<td>Schizophrenia, undifferentiated type</td>
<td>4</td>
</tr>
<tr>
<td>Schizo-affective disorder</td>
<td>2</td>
</tr>
<tr>
<td>Schizophreniform disorder</td>
<td>1</td>
</tr>
<tr>
<td>Psychotic disorder not otherwise specified</td>
<td>1</td>
</tr>
<tr>
<td>Secondary DSM-IV diagnosis</td>
<td>1</td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td>1</td>
</tr>
<tr>
<td>Schizotypal personality disorder</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Duration of illness (weeks [median=26])</td>
<td>183.0</td>
</tr>
<tr>
<td>BPRS total score</td>
<td>51.5</td>
</tr>
<tr>
<td>GAF score (past month)</td>
<td>29.5</td>
</tr>
<tr>
<td>Comprehensive Assessment of Symptoms and History</td>
<td>7</td>
</tr>
<tr>
<td>Delusions</td>
<td>3.8</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>2.1</td>
</tr>
<tr>
<td>Bizarre behavior</td>
<td>2.2</td>
</tr>
<tr>
<td>Formal thought disorder</td>
<td>1.3</td>
</tr>
<tr>
<td>Catatonia</td>
<td>0.4</td>
</tr>
<tr>
<td>Alogia</td>
<td>1.4</td>
</tr>
<tr>
<td>Affective flattening</td>
<td>1.2</td>
</tr>
<tr>
<td>Inappropriate affect</td>
<td>1.1</td>
</tr>
<tr>
<td>Avolition/apathy</td>
<td>2.3</td>
</tr>
<tr>
<td>Anhedonia/Asociality</td>
<td>2.6</td>
</tr>
<tr>
<td>Attentional impairment</td>
<td>1.3</td>
</tr>
<tr>
<td>Abnormal Involuntary Movement Scale score</td>
<td>1.5</td>
</tr>
</tbody>
</table>

PET Scans

FDG was administered intravenously (mean=7.8 mCi, SD=0.6). Emission scanning commenced 40 minutes following the injection of FDG. The data were acquired with an ECAT EXACT HR+ PET scanner in three-dimensional mode (17). Subjects were positioned in the PET camera with the canthomeatal line parallel to the in-plane field of view. Before acquisition of the FDG data, a 68Ge/68Ga transmission scan was acquired for 5 minutes to correct for the attenuation of radiation. A single emission scan of 20 minutes’ duration was acquired. The data were reconstructed by using ECAT version 7.2 software implementation of filtered back-projection (4-mm Hanning filter) with a pixel size of 1.8x1.8x2.4 mm. The average interval between FDG injection and commencement of PET scanning was 45.04 minutes (SD=1.51), with no difference between healthy subjects (mean=44.64 minutes [SD=1.39, range=40–46]) and schizophrenia patients (mean=45.55 minutes [SD=1.57, range=44–50]).

MRI Scans

T1-weighted axial MRI scans were acquired with the GE Signa 5x system (General Electric, Milwaukee) (TR=24 msec, TE=5 msec, flip angle=40°, slice thickness=1.2 mm, pixel matrix=256x256, field of view=23 cm, total slices=128). The interval between FDG PET and MRI scans averaged approximately 1 day for schizophrenia patients (mean=0.92 days, range=0–3) and 2 weeks for comparison subjects (mean=13.6 days, range=3–70).

Uptake Condition

During the 40-minute FDG uptake period, all subjects carried out a visual attention task that required separation of target stimuli from competing surroundings as in our earlier PET study (12). Subjects visually fixated upon a dot corresponding to the center of the screen then looked at stimuli positioned horizontally at 2° to the right or to the left of the central fixation point. The target stimulus (the letter O) either appeared alone as an upper-case character or as a lower-case character surrounded by eight other letters (Figure 1). In half of the trials, the letter C or the digit zero (0) were presented as distracters. The subject’s task was to click on the mouse each time he detected the letter O either alone or surrounded by small letters, ignoring the C and the 0, and to press on the right button for a right-sided target and the left button for a left-sided target. The overall size of the stimulus was controlled so that the big letters were of the same dimensions as the pattern of small letters surrounded by flankers (i.e., each stimulus display was 19×22 mm). Each display was flashed for 150 msec. There were four experimental runs, each 264 seconds in duration, with a brief rest interval of 24 seconds between each session. Each run began with a 24-second period of blank screen followed by a block of 12 display stimuli, eight of one type and four drawn at

TABLE 1. Clinical Characteristics of a Group of Never Medicated Patients With Schizophrenia Recruited for a PET Study of Regional Glucose Metabolism in the Fronto-Striato-Thalamic Circuit

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Schizophrenia Patients (N=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary DSM-IV diagnosis</td>
<td>N</td>
</tr>
<tr>
<td>Schizophrenia, paranoid type</td>
<td>4</td>
</tr>
<tr>
<td>Schizophrenia, undifferentiated type</td>
<td>4</td>
</tr>
<tr>
<td>Schizo-affective disorder</td>
<td>2</td>
</tr>
<tr>
<td>Schizophreniform disorder</td>
<td>1</td>
</tr>
<tr>
<td>Psychotic disorder not otherwise specified</td>
<td>1</td>
</tr>
<tr>
<td>Secondary DSM-IV diagnosis</td>
<td>1</td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td>1</td>
</tr>
<tr>
<td>Schizotypal personality disorder</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Duration of illness (weeks [median=26])</td>
<td>183.0</td>
</tr>
<tr>
<td>BPRS total score</td>
<td>51.5</td>
</tr>
<tr>
<td>GAF score (past month)</td>
<td>29.5</td>
</tr>
<tr>
<td>Comprehensive Assessment of Symptoms and History</td>
<td>7</td>
</tr>
<tr>
<td>Delusions</td>
<td>3.8</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>2.1</td>
</tr>
<tr>
<td>Bizarre behavior</td>
<td>2.2</td>
</tr>
<tr>
<td>Formal thought disorder</td>
<td>1.3</td>
</tr>
<tr>
<td>Catatonia</td>
<td>0.4</td>
</tr>
<tr>
<td>Alogia</td>
<td>1.4</td>
</tr>
<tr>
<td>Affective flattening</td>
<td>1.2</td>
</tr>
<tr>
<td>Inappropriate affect</td>
<td>1.1</td>
</tr>
<tr>
<td>Avolition/apathy</td>
<td>2.3</td>
</tr>
<tr>
<td>Anhedonia/Asociality</td>
<td>2.6</td>
</tr>
<tr>
<td>Attentional impairment</td>
<td>1.3</td>
</tr>
<tr>
<td>Abnormal Involuntary Movement Scale score</td>
<td>1.5</td>
</tr>
</tbody>
</table>

a Two subjects initially diagnosed as schizophreniform were determined at follow-up to have schizophrenia.
b Symptoms had fully remitted within 6 months but patient continued to receive antipsychotic medication.
c Subject had consumed no alcohol in the week preceding the study evaluations; no evidence of alcohol withdrawal syndrome or clinical instability was noted as a result of alcohol abstinence; criteria for alcohol dependence were not met.
d Dimensions rated on a 0–5 scale, with 0=none, 1=questionable, 2=mild, 3=moderate, 4=marked, and 5=severe.
e For the two patients with scores >0.

random from the other three types to maintain expectancy. The four display types were large letter in 1) left or 2) right hemifield and small letter with flankers in 3) left or 4) right hemifield. The order of runs was counterbalanced across subjects. The subjects were signaled with a blue flash in case of a miss or an error. Subjects were trained before FDG administration.

Data Analysis

Significance probability maps of the datasets were created using two complementary methods, statistical parametric mapping software (SPM) (18) and the FMRIB Software Library (FSL), version 5.00 (www.fmrib.ox.ac.uk/fsl). We chose SPM2 to survey the entire smoothed brain and FSL with our own R programs to evaluate the primary brain region, the thalamus with unsmoothed MRI template coregistered images. For SPM2, the FDG images were first spatially normalized to the FDG template defined with the SPM 99 software (and subsequently converted for SPM2) using the default parameters for spatial normalization, adjustment for whole brain metabolic rate by covariance, and 8-mm smoothing. Both SPM and FSL data are unitless. The normalized images were then closely inspected to ensure proper alignment. A schizophrenia and healthy group comparison was made by using the two-sample t test criterion. The t test results were then displayed on the FDG template. For closer examination of the thalamus, we used FSL processing. MRI anatomical images had the brain extracted from the skull with the Brain Extraction Tool (19) and were placed in standard position (anterior and posterior commissures in the same plane, midline x=0, vertical level z=0) using a 6-parameter transform. No images were smoothed by filtering using the FSL tool. The FDG images were coregistered to each individual’s anatomical image with the 6-parameter transform and the FMRIB Linear Image Registration Tool (20). Next, each person’s MRI anatomical images were transformed to Montreal Neurological Institute (MNI) coordinates with a 12-point transformation and the transformation matrix used to similarly transform the unsmoothed FDG images, which were then divided by the whole brain average metabolic rate. The FDG images then had pixel-by-pixel t tests computed using our own program in the R computer language (http://www.r-project.org). Since we had already reported medial dorsal nucleus FDG decreases in two earlier and entirely independent samples (11), both predicting the finding in advance of SPM application, images were thresholded at p<0.05. This provides an alternate coregistration method unbiased by systematic group differences in brain rGMR, since the FDG itself did not enter the coregistration transformations. The FDG data was unsmoothed (in the FSL analysis) for maximum spatial resolution in a priori hypothesis testing. While SPM maps are typically corrected for multiple comparisons because of their exploratory nature, this analysis was confirmatory of our earlier FDG studies and neuroanatomical studies of the medial dorsal nucleus.

We examined the quality of the FMRIB Linear Image Registration Tool coregistration and standardization to MNI space on the final MNI FDG and anatomical images. We chose the horizontal slice at z=12 because it was close to the center of the medial dorsal nucleus. Three regions were chosen as internal fiducial marks: 1) along the x=0 midline, the transition from the gray matter of the posterior cingulate (Brodman’s areas 29, 30) to the white matter of the corpus callosum; 2) along the x=–4 anteroposterior line, the transition from the gray matter of areas 24 and 32 to the white matter or the genu of the corpus callosum; and 3) the fluid space between the two halves of the posterior thalamus along the y=–24 line. These landmarks were chosen because 1) there would be a rapid change in values along a unidimensional column of pixels for both PET (from high values in gray to low values in white) and for MRI (from low values in gray to high values in white) and 2) they were adjacent to the medial dorsal nucleus. The medial dorsal nucleus is more difficult to trace, and intertracer variability would also have been involved in its use as an internal landmark. Both PET and MRI would be expected to have low values in the midline position chosen. The PET and MRI values were differentiated (x_n–xn+1) and the location of the maximum value determined. This location should be the same for every MNI standardized brain and appear at the same location in every MNI standardized PET. Variation in the quality of the 12-point affine standardization and its success in aligning all structures can be assessed by examining the MNI MRI data. The MNI images were examined in MRicro, and an R program created in house by one of the authors (M.S.B.) was used to calculate the first derivative strips. The MNI pixel size was 2 mm. The posterior cingulate landmark had a mean y=–46.5, and 85% of subjects fell from –45 to –49, an error of one discrete pixel. The Talairach atlas shows this location at y=–44. The midline was found on the MRI at x=0 in 62% and an error of one pixel on either side (–2 or 2) in 92%. The midline was found on PET at x=0 in 38% and within an error of one pixel on either side in 100%. The differences in distance between the location of maximum first derivative points on MRI and PET were 0 pixels for 57% of subjects and within one pixel error for 71%. The anterior cingulate landmark was somewhat more variable anatomically with a mean of 26.7, 71% of subjects within the one pixel error, and the Talairach atlas position y=30. The differences between the location of maximum first derivative points between MRI and FDG were 0 for 53% and within one pixel for 85%. Thus, the combined error of brain standardization and coregistration (PET locations) was within one pixel (2 mm) in 80–90% of landmark examinations and adequate for regional examination of the thalamus. This compares favorably with the typical width of the median dorsal nucleus, 6–8 mm. For comparison, the thalamus is 21 mm wide (x) and 30 mm long (y) at this level; a medial dorsal nucleus bounding box in the Talairach atlas at this level (z=12) would be 9 mm in the x dimension and 13 mm in the y dimension.

The thalamus and medial dorsal nucleus were also assessed using a stereotaxic region of interest atlas approach for confirmation (21). This method has stored coordinates for MNI brain structures including the thalamus and the medial dorsal nucleus. Last, one of the authors (E.M.K.) traced the whole thalamus, medial dorsal nucleus, and pulvinar on raw unresliced MRI templates exactly as previously reported (4, 5). FDG images were coregistered to these anatomical images as before. Relative meta-
Results

SPM Mapping

Patients with schizophrenia had lower relative metabolic rates in the thalamus and lateral prefrontal cortex, especially at the orbitofrontal levels (Figure 2). A significant reduction in metabolic rate was also noted in additional cortical areas, including the insula, temporal lobes, parietal lobe (Brodmann's area 39), and the anterior cingulate at its most dorsal level (Brodmann's area 24) (22). The posterior putamen was also less active in ill subjects. The centers of the thalamic regions were at -6, -16, 8 and at 6, -16, 8 (total volume=297 voxels, $z_{max}$=2.37 and 2.32, respectively).

Pixel-by-Pixel Examination of the Thalamus

Coregistration is illustrated in Figure 3, and detailed examination of Talairach z=12 slice in Figure 4. The red area in Figure 4 indicates pixels where the healthy subjects have higher relative metabolic rates than patients with schizophrenia (p<0.05, two-tailed), which extends across the region of the medial dorsal nucleus and pulvinar.

Stereotaxic Region-of-Interest Examination of the Thalamus

The medial dorsal nucleus had a relative metabolic rate significantly higher in healthy volunteers (mean=11.8, SD=0.64) than patients with schizophrenia (mean=11.2, SD=0.64) ($t=2.88, df=23, p=0.008$). The whole thalamus did not differ between normal subjects and patients (mean=10.0 [SD=0.43] and 9.7 [SD=0.49], respectively). Repeated measures two-way analysis of variance (ANOVA) with whole thalamus and medial dorsal nucleus for the right and left hemisphere revealed a significant diagnostic group difference ($F=6.66, df=1, 23, p<0.02$) and a significant interaction between group and region of interest (whole thalamus, medial dorsal nucleus alone) ($F=6.45, df=1, 23, p<0.02$), indicating that the metabolic group difference effect for the medial dorsal nucleus was significantly greater than for the whole thalamus. There was no significant diagnostic group-by-hemisphere or diagnostic group-by-hemisphere-by-structure interaction.

MRI Template-Traced Nuclei

A lower metabolic rate in the medial dorsal nucleus for patients with schizophrenia (mean=2.59, SD=0.29) versus healthy subjects (mean=2.76, SD=0.28) was confirmed ($F=4.62, df=1, 18, p<0.05$), and this was specific to the medial dorsal region when compared with the remainder of the thalamus in a three-way ANOVA (diagnostic group [healthy, schizophrenia] by structure [medial dorsal, remainder of thalamus] by hemisphere [right, left]), with a significant group-by-structure interaction ($F=5.51, df=1$, $p<0.05$).
It is of interest that all subjects except one schizophrenia patient had higher relative values in the medial dorsal nucleus than in the remainder of the thalamus. While the relative metabolic rate was lower in the pulvinar in patients with schizophrenia, this difference did not reach statistical significance. The correlation between the stereotaxic medial dorsal nucleus and the traced medial dorsal nucleus was 0.51.

**Task Performance**

All healthy subjects and all but one patient performed the task above chance levels and completed responses in every trial block during the uptake period. Removal of the one schizophrenia subject who did not perform the task did not alter the significance of the group comparison t test on the medial dorsal nucleus (t=2.31, df=22, p=0.03) or the diagnostic group difference with ANOVA (F=4.32, df=1, 22, p<0.05).

**Discussion**

Our results confirm earlier studies that reported relative metabolic rate reduction in the thalamus and lateral prefrontal cortex in patients with schizophrenia, extending those findings in several important ways. We have demonstrated these findings in a group of never medicated patients, indicating that these reductions could not be due to medication artifacts. Further, the use of an uptake condition demonstrated to have thalamic activation in earlier PET studies enhances the strength of the conclusions. Replicating the results with methodologically complementary pixel-by-pixel t test methods and traced templates allows both whole-brain exploration and the specific regions to be exhaustively tested. These specific methodological steps appear to have enhanced the power of the FDG PET images to confirm group differences in comparison with some earlier studies that used medicated patients, an eyes-closed resting condition, or lower-resolution image acquisition.

The limitations of our present study include the relatively modest number of never medicated subjects, image resolution and coregistration quality, and the use of relative metabolic rates. While somewhat smaller than two earlier studies of never medicated patients (6, 9), we had adequate power to detect differences in thalamic rGMR. Methods are currently being explored to overcome image resolution limitations by applying a partial volume correction to the image data. However, this correction has not been fully validated and has the potential to introduce a large bias into the analysis (23, 24). With a structure as small as the thalamus or medial dorsal nucleus (6–9 mm wide), image coregistration and the quality of structure position standardization are critical. We evaluated joint error in our examination of FDG image coordinates and found that about 75% of subjects’ FDG landmarks adjacent to the medial dorsal nucleus lay within one 2-mm pixel of a standard location and of the matching MRI template. Last, we thresholded the significance probability images at the p<0.05 level without correction. This lacks correction for multiple t testing, but since the study was focused specifically on the medial dorsal region (the relative metabolic rate of which was previously reported to be reduced [11]), and the dorsolateral prefrontal cortex was previously explored with significance probability mapping (25), it seemed appropriate to consider the maps as confirmatory rather than exploratory. The medial dorsal region selected a priori and assessed using stereotaxic position also yielded a p=0.0083 probability (0.0041 one-tailed in replication). While absolute metabolic rates for the medial dorsal nucleus could have been obtained with the FDG method, there is no evidence that total brain metabolic rate differs between normal subjects and patients with schizophrenia. Thus the small regions of thalamic and prefrontal metabolism reductions seen in Figure 2 seem unlikely to have artifactually arisen on the basis of other large areas of cortical metabolic rate shifts.

One might argue that patients with schizophrenia lacked motivation to perform the task and that differences in the
Our use of unmedicated and previously untreated patients is important in relating the findings to schizophrenia rather than medication effects. Studies have reported decreases in relative metabolic rate in the frontal lobe with clozapine (27) and haloperidol (28) but not with risperidone (29), so the use of unmedicated patients is especially important for establishing this finding. It is interesting that patients with higher dorsolateral prefrontal metabolic rates were reported likely to respond to clozapine (30).

Other studies have also confirmed reduced thalamic FDG uptake in schizophrenia (31, 32) (reviewed elsewhere [11]). Last, using functional MRI (fMRI), reduced activation in the thalamus has been observed. The anterior thalamic region as well as the medial and anterior frontal lobe had reduced fMRI blood oxygenation level-dependent activation in medicated patients with schizophrenia during the continuous performance test of vigilance (33), both areas adjacent to but not exactly overlapping with our regions. Less thalamic activation in schizophrenia was also seen in other fMRI studies (34). However, significant thalamic activation was observed in medicated patients with schizophrenia but not in normal subjects during performance of the Sternberg memory task (35), although apparently no direct map of group differences in activation was presented.

Examination of the pixel-by-pixel maps indicates that medial and posterior portions of the thalamus appear to be the main regions of the thalamus with decreased metabolic rates. The thalamus comprises multiple nuclei that relay and filter sensory and higher-order inputs to and from the cerebral cortex and limbic structures (3). The biggest thalamic substructures, the medial dorsal nucleus and the pulvinar (both visible on MRI), are of particular interest in attention because of their reciprocal connections with the prefrontal and temporal regions and because their size and metabolic rate appear diminished in imaging studies of patients with schizophrenia. The medial dorsal nucleus has strong interconnections with the dorsolateral prefrontal cortex (36), and the connections of the medial dorsal nucleus have been used to define the prefrontal cortex (37), a key area of executive action and focusing of attention thought to be defective in schizophrenia (reviewed by Buchsbaum and Hazlett [2]). Crosson (38) suggests the medial dorsal nucleus as a critical element in an attentional “selective engagement” system that impacts semantic functions in schizophrenia.

It should be noted that the use of never medicated patients may be especially important in establishing regional metabolic differences in the thalamus and that even thalamic volume may be affected by neuroleptics (39). The recent report of decreased thalamic D_{2} receptor binding in the medial portion of the thalamus in drug-naïve patients with schizophrenia (40) indicates that both chronic and acute medication effects could influence thalamic metabolic rate. That important study not only supports a role for the medial thalamus in schizophrenia but also sug-
gests that the medial regions of the thalamus are potentially informative for understanding medication effects as well.

Until the mid-1970s it was thought that the prefrontal cortex received its thalamic input solely from the medial dorsal nucleus; however, it is now apparent that the pulvinar also contributes to its innervation (41, 42). The pulvinar, important in visual and possibly auditory attention (43–45), also has prominent interconnections with the parietal and temporal lobe. The current findings of decreased metabolic rate in the medial and posterior regions of the thalamus, prefrontal cortex, and posterior temporal and parietal regions tend to confirm a network of frontothalamic and parieto-temporal-thalamic rMRI concurrently diminished in patients carrying out a task demonstrated to activate the pulvinar (12).

Much work in the field of visual neurophysiology suggests that the pulvinar signals the importance or relevance of stimuli that fall inside classically defined visual receptive fields (43). Neurophysiological studies in animals have found that neurons of the pulvinar respond specifically to stimuli to be the target of a saccade (46). These findings suggest that the pulvinar may play a role in “visual salience,” that is, the attention-related selection of a target (44), and “directed attention” (12). The anatomy and projections of the pulvinar suggest functional cortical-thalamic-cortical loops involved in a variety of functions including salience, attention, and working memory (43). The fact that both the medial dorsal nucleus and the pulvinar express D2 type dopamine receptors (40), together with their demonstrated role in selective attention, fits well with recent formulations of schizophrenia as a hyperdopaminergic state that leads to an aberrant assignment of salience to elements of one’s experience (47). According to that model, both typical and atypical antipsychotics “dampen the salience” of abnormal experiences via blockade of D2 receptors and by doing so, allow the resolution of symptoms.

References


3. Jones E: Cortical development and thalamic pathology in schizophrenia. Schizophr Bull 1997; 23:483–501

4. Byne W, Buchsbaum MS, Kemether E, Hazlett EA, Shihwari A, Mitropoulou V, Siever LJ: Magnetic resonance imaging of the thalamic mediiodorsal nucleus and pulvinar in schizophrenia and schizotypal personality disorder. Arch Gen Psychiatry 2001; 58:133–140


Received March 25, 2004; revision received June 7, 2004; accepted June 14, 2004. From the Wallace-Kettering Neuroscience Institute. Address correspondence and reprint requests to Dr. Lehrer, Wright State University Department of Psychiatry, c/o The Wallace-Kettering Neuroscience Institute, 3533 Southern Blvd., Suite 5200, Kettering, OH 45429; Doug.Lehrer@kmcnetwork.org (e-mail).

Major support was provided by Wallace-Kettering Neuroscience Institute. Use of imaging resources supported by the United States Air Force, Air Force Research Laboratory (AFRL/HEOP), Air Force Materiel Command, under cooperative agreement F33615-98-2-6002.

The authors thank Tonya Perkins, Marilyn Brackney, Kelly Dunigan, Kerry Kovacs, Yuliya Lelchuk, and Drs. Jogesh Mukherjee, Martin Satter, and Kingwai Chu for technical assistance. The authors also thank Drs. Bing Shi, T.K. Narayanan, and Steve Mattmuller for radiopharmaceutical production.
THALAMIC AND PREFRONTAL METABOLISM IN SCHIZOPHRENIA

Evidence for Onset of Antipsychotic Effects Within the First 24 Hours of Treatment

Shitij Kapur, M.D., Ph.D.
Tamara Arenovich, M.S.
Ofer Agid, M.D.
Robert Zipursky, M.D.
Stacy Lindborg, Ph.D.
Barry Jones, M.D.

Objective: It is widely held that there is a delayed onset of antipsychotic action and that any early effects represent nonspecific behavioral effects. Recent research has shown that antipsychotic action begins within the first week. The authors tested the hypothesis that psychosis improves within the first 24 hours of antipsychotic treatment.

Method: In this multicenter, double-blind, placebo-controlled study, 311 patients with a diagnosis of schizophrenia spectrum disorder and an acute exacerbation were randomly assigned to receive 10 mg i.m. of olanzapine, 7.5 mg i.m. of haloperidol, or intramuscular placebo. Subjects were rated with structured rating scales (Positive and Negative Syndrome Scale and Clinical Global Impression) at baseline, 2 hours, and 24 hours.

Results: The olanzapine and haloperidol groups showed greater resolution of overall symptoms than the placebo group; for the olanzapine group, this effect was evident at 2 hours. A factor analysis showed that an independent change in psychosis (which included conceptual disorganization, hallucinatory behavior, unusual thought content) was evident within the first 24 hours for both drugs. This improvement in core psychosis was not mediated unidirectionally by changes in nonspecific behavioral effects or other psychopathology.

Conclusions: These data suggest that the onset of antipsychotic action is early and that the magnitude of this action grows with time. This clinical reality calls into question some prevailing hypotheses regarding the mechanism of action of antipsychotics and suggests that antipsychotic action may be more proximally related to the blockade of dopamine transmission than was originally thought.

Antipsychotics were introduced to modern psychiatric treatment more than half a century ago, yet their mechanism of action remains an issue of active debate (1). A central question regarding antipsychotic action is the speed of “onset” of antipsychotic response. It is widely held that there is a “delayed onset” of antipsychotic response, somewhere in the range of 2–3 weeks, an idea now embedded in standard psychiatric texts (2, 3). The earlier effects are thought to be a reflection of the nonspecific behavioral effects of antipsychotics on aspects such as agitation, excitement, and uncooperativeness. However, a recent meta-analysis by Agid et al. (4), which included data from 7,450 patients in 42 double-blind, active-drug or placebo-controlled trials, did not support this “delayed onset” hypothesis. The authors found that 1) there is a definite change in psychosis by the end of the first week, 2) the degree of improvement in the first week exceeds that in each subsequent week, and 3) this improvement is over and above any change accorded to placebo or nonspecific behavioral improvement (4). If antipsychotic response is established within the first week, how early can it be apparent?

When antipsychotics were introduced in the 1950s and 1960s, it was routinely noted that they were effective within the first few days (5, 6). In the 1970s the interest in “rapid neuroleptization” (7) led to several studies that compared the safety and efficacy of different doses, parenteral versus oral administration, and low-potency versus high-potency agents in the first few days of treatment (8). However, 1) many of these reports included small and uncontrolled case series (7); 2) they did not include a placebo control (7, 9–15); 3) raters were not blind to treatment assignment (7, 14); 4) the response measures reported often did not differentiate between effects on psychosis and nonspecific effects on behavior (10, 13); and 5) where the two measures were separately reported, both were shown to improve, and none of these studies tested whether the improvement in psychosis is distinct from, or secondary to, a nonspecific behavioral effect (9, 11, 12, 14–16). These studies all pointed to a rapid response—however, they suffered from two major confounding factors. First, most of these studies lacked a placebo control. The patients included in studies of acute exacerbation were often at the apogee of their worsening, and investigators could have had a bias to overestimate this worsening to include them in the trial. Both of these factors may lead to the appearance of a rapid response—an issue that can only be resolved with the inclusion of a placebo-treated control group. Second, these studies did not include appropriate subscales or did not carry out the required analyses that permit one to determine whether
EARLY ONSET OF ANTIPSYCHOTIC EFFECTS

TABLE 1. Correlation Matrix of Percentage Reduction in Baseline Scores on Brief Psychiatric Rating Scale (BPRS) Items at 24 Hours After Treatment Onset in Patients With Acute Exacerbation of Psychosis Who Received Intramuscular Olanzapine (N=113) or Haloperidol (N=112)a

<table>
<thead>
<tr>
<th>BPRS Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Conceptual disorganization</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Hallucinatory behavior</td>
<td>0.28**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Excitement</td>
<td>0.20*</td>
<td>0.22*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Grandiosity</td>
<td>0.12</td>
<td>0.06</td>
<td>0.08</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Suspiciousness/persecution</td>
<td>0.25*</td>
<td>0.08</td>
<td>0.24*</td>
<td>0.11</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Hostility</td>
<td>0.17</td>
<td>0.15</td>
<td>0.53**</td>
<td>0.05</td>
<td>0.24*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Blunted affect</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.03</td>
<td>0.13</td>
<td>0.01</td>
<td>0.02</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Emotional withdrawal</td>
<td>0.11</td>
<td>0.23*</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.03</td>
<td>0.08</td>
<td>0.13</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Somatic concerns</td>
<td>0.12</td>
<td>0.15</td>
<td>0.17*</td>
<td>0.24*</td>
<td>0.22*</td>
<td>0.09</td>
<td>0.18*</td>
<td>0.14</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Anxiety</td>
<td>0.09</td>
<td>0.19*</td>
<td>0.41**</td>
<td>0.02</td>
<td>0.11</td>
<td>0.27**</td>
<td>-0.02</td>
<td>0.13</td>
<td>0.16</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>11. Guilt feelings</td>
<td>0.23*</td>
<td>0.09</td>
<td>0.02</td>
<td>0.06</td>
<td>0.02</td>
<td>0.0</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.08</td>
<td>0.15</td>
<td>1.00</td>
</tr>
<tr>
<td>12. Tension</td>
<td>0.13</td>
<td>0.23*</td>
<td>0.56**</td>
<td>0.11</td>
<td>0.14</td>
<td>0.52**</td>
<td>0.03</td>
<td>0.0</td>
<td>0.18*</td>
<td>0.50**</td>
<td>0.11</td>
</tr>
<tr>
<td>13. Mannerisms and posturing</td>
<td>0.15</td>
<td>0.13</td>
<td>0.08</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.03</td>
<td>0.04</td>
<td>-0.07</td>
<td>0.12</td>
<td>0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>14. Depression</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.04</td>
<td>-0.05</td>
<td>0.10</td>
<td>0.07</td>
<td>0.14</td>
<td>0.15</td>
<td>0.27**</td>
<td>0.17</td>
<td>0.31**</td>
</tr>
<tr>
<td>15. Motor retardation</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.06</td>
<td>-0.08</td>
<td>0.01</td>
<td>0.05</td>
<td>0.14</td>
<td>0.09</td>
<td>0.18*</td>
<td>-0.08</td>
<td>0</td>
</tr>
<tr>
<td>16. Uncooperativeness</td>
<td>0.20*</td>
<td>0.06</td>
<td>0.46**</td>
<td>0.05</td>
<td>0.28**</td>
<td>0.61**</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.10</td>
<td>0.18*</td>
<td>0.09</td>
</tr>
<tr>
<td>17. Unusual thought content</td>
<td>0.17</td>
<td>0.35**</td>
<td>0.12</td>
<td>0.02</td>
<td>0.10</td>
<td>0.09</td>
<td>0.04</td>
<td>-0.05</td>
<td>0.11</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>18. Disorientation</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.05</td>
<td>-0.05</td>
<td>-0.07</td>
<td>0.09</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.12</td>
<td>0.10</td>
<td>0.24*</td>
</tr>
</tbody>
</table>

a Patients with a clinical diagnosis of schizophrenia, schizophreniform disorder, or schizoaffective disorder received one intramuscular dose of 10 mg of olanzapine or 7.5 mg of haloperidol within a 24-hour treatment period (with optional second and third injections given 2 or more and 4 or more hours after the first and optional second injections, respectively).

b Items with correlation coefficients >0.3 (highlighted in boldface type) were included in factor analysis.

*p<0.01. **p<0.001.

this early response in psychosis is primary or secondary to the effects of the drug on agitation and excitement.

A very good opportunity to test this question was provided by a recent large-scale clinical trial that compared the efficacy of haloperidol, olanzapine, and placebo in patients with schizophrenia during the first 24 hours of treatment (17). Unlike the previous reports, this study 1) was large (N=311), 2) was placebo-controlled, 3) had raters who were blind to drug assignment, 4) used standardized and comprehensive rating scales (Positive and Negative Syndrome Scale and Brief Psychiatric Rating Scale) with separable items for psychosis and agitation, and 5) obtained data at 2 hours and then at 24 hours after the onset of treatment. Using these data we investigated the following questions: Is there an improvement in core psychotic symptoms in the first 24 hours? Is this effect clearly different from the effect of placebo? Is this early improvement in psychosis a secondary consequence of nonspecific tranquillizing effect of the drugs?

Method

Study Subjects

The study was an Eli Lilly–sponsored, multisite, international study comparing two active drugs (haloperidol and olanzapine) in their intramuscular form versus placebo in a double-blind, randomized, controlled trial. Local ethics review boards approved the study protocol, including the use of placebo, given the hospitalized status of all participating patients, the 2:2:1 randomization ratio for active treatment versus placebo, the brief duration of the study (24 hours), and the use of active medication based on the clinical judgment of the investigator at the time of randomization. Written informed consent was obtained from all patients and from a relative or legal representative when required by local law or custom.

Recently hospitalized patients ages 18 years or older who had been assessed by the study investigators and given a clinical diagnosis of DSM-IV schizophrenia, schizophreniform disorder, or schizoaffective disorder and who exhibited an episode of acute agitation in the context of psychosis were included. Patients were included in the study if they demonstrated a total score of 14 or higher (of a maximum of 35) on the Positive and Negative Syndrome Scale excitement component (which included items measuring tension, uncooperativeness, hostility, poor impulse control, and excitement) with at least one item score >4. Patients with significant, unstable, medical disorders and those who were too agitated to provide informed consent or to cooperate with the requirements of the study were not included in this trial.

Clinical Design and Outcome Measures

The study consisted of a screening period and a 24-hour intramuscular treatment period. Patients were not allowed to receive any antipsychotic treatment during the screening period, which lasted for a minimum of 2 hours. On entering the treatment period, patients were randomly allocated to treatment with 10 mg of olanzapine or 7.5 mg of haloperidol within a 24-hour treatment period (with optional second and third injections given 2 or more and 4 or more hours after the first and optional second injections, respectively). Benzodiazepines were allowed at the discretion of the treating clinician, but only for patients who received more than one injection of active treatment/placebo. Those who received two injections of drug/placebo could receive a dose of a benzodiazepine (2 mg of lorazepam or equivalent) no earlier than 1 hour after the injection. Those who received a third injection of the drug/placebo could get the second dose of benzodiazepine at least 1 hour after that injection.
The main intent of this study was to determine whether there was drug-induced early (within first 24 hours) improvement in psychosis and how this effect was related to drug-induced changes in nonspecific behavioral symptoms. Although the BPRS scale has been subjected to factor analysis, most of these analyses were derived from large-scale studies of stable patients, and it is now clear that the factor structure is not invariant over time (19). Because our primary interest was in the pattern of change in the first 24 hours, we undertook a factor analysis of the change scores on all the rated items of the BPRS to get an empirically valid estimate of the factors of early antipsychotic response. To relate this factor-analysis-derived answer to more conventional measures, we also asked whether the improvement in the conventional BPRS factor representing psychosis (items: conceptual disorganization, hallucinatory behavior, grandiosity, unusual thought content) was explained by changes in the conventional BPRS factors representing activation and hostility (items: excitement, tension, mannerisms, suspiciousness, hostility, un cooperativeness). Further, because the five items that constitute the Positive and Negative Syndrome Scale excitement component are also markers of nonspecific behavioral effects and have shown to improve with drug treatment in 24 hours, we examined if this factor could explain this change in the BPRS psychosis factor. Finally, we used the mediator analysis strategy of Kraemer et al. (20) to examine if these early changes in core psychosis could be explained by changes in any of the nonspecific factors. Statistical testing, as described in the following sections, was performed with the SAS System version 8.2 (SAS, Cary, N.C.). An alpha of 0.05 for significance was used.

### Determining the Factors of Early Response

The percentage of reduction in baseline score was calculated for each Positive and Negative Syndrome Scale–derived BPRS item at 2 hours and 24 hours. The following equation was used to calculate these changes:

$$\% \text{change} = \left( \frac{x_{0} - x_{j}}{x_{0}} \right) \times 100\%$$

where $y_{ij}$ is the change in baseline score for BPRS item $i$ at time $j$; $x_{ij}$ is the raw score for BPRS item $i$ at time $j$; $x_{0}$ is the baseline score for BPRS item $i$; $i$ is the BPRS item, ranging from 1–18; and $j$ is the time since onset of study, equal to 2 hours or 24 hours.

To understand the nature of the relationship between these changes at a given point in time, we developed a correlation matrix of the 24-hour change scores (Table 1), and the data were subjected to a series of factor analyses. Because we were interested in identifying factors in drug-induced olanzapine or haloperidol change at this stage, the first factor analyses were based on the data from patients who received active treatment. Subsequent comparisons included the data from patients who received placebo.

An unweighted least-squares approach was implemented, and factor solutions were then rotated using a varimax (orthogonal) rotation. Initially, all 18 items were incorporated into the analysis, and those items that did not contribute meaningfully to any of the factors (i.e., had a maximal loading less than 0.30) were removed from the analysis. This step was done to conserve statistical power and to make the analysis and its interpretation as clean and simple as possible.

### Drugs Versus Placebo Effects on Factor Scores

Factor scores were calculated for each patient by adding up their change scores for each of the items that loaded on a corresponding factor. Some items loaded into more than one factor (e.g., excitement loaded on both factor 1 and factor 2), in which case they were added to both factor scores. Differences in the degree of improvement in these factors with drug versus placebo were compared at the 2 and 24 hour time points by using analysis of variance. If the overall model was significant, post hoc $t$ tests were done to ascertain the statistical significance of differences at given time points. A Bonferroni adjustment, relative to the number of multiple comparisons in each model, was applied to the $p$ values as a conservative way to reduce the risk of finding false positive results and to maintain an overall alpha level of 0.05.

### Changes in Other Factors

To test whether changes in other factors accounted for the improvement in psychosis, a series of mediator analyses and analyses of covariance (ANCOVAs) was performed. The definitions of mediator/moderator used in these analyses are those defined by Kraemer et al. (20). By definition, change scores cannot be moderators because they are not baseline characteristics that occur independent of and prior to treatment. Thus, we were left with only the question of whether any of the factor scores were mediating the other factors.

The method proposed by Kraemer et al. (20) involves the construction of three equations:

1. $\text{mediator} = \beta_0 + \beta_1 \text{ treatment}$
2. $\text{outcome} = \beta_2 + \beta_3 \text{ treatment}$

---

**Table 1.**

<table>
<thead>
<tr>
<th></th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08</td>
<td>0.08</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16</td>
<td>0.09</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.03</td>
<td>0.07</td>
<td>0.04</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.37**</td>
<td>0.07</td>
<td>0.04</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.20*</td>
<td>0.13</td>
<td>0.07</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.12</td>
<td>0.07</td>
<td>0.19*</td>
<td>0.03</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 2. Factor Analysis of Percentage Reduction in Baseline Scores on Brief Psychiatric Rating Scale (BPRS) Items at 24 Hours After Treatment Onset in Patients With Acute Exacerbation of Psychosis Who Received Intramuscular Olanzapine (N=113) or Haloperidol (N=112)

<table>
<thead>
<tr>
<th>BPRS Item</th>
<th>Factor 1: Agitation/Excitement</th>
<th>Factor 2: Hostility</th>
<th>Factor 3: Psychosis</th>
<th>Factor 4: Other Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tension</td>
<td>79</td>
<td>29</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Anxiety</td>
<td>59</td>
<td>7</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Excitement</td>
<td>58</td>
<td>47</td>
<td>17</td>
<td>-10</td>
</tr>
<tr>
<td>Uncooperativeness</td>
<td>20</td>
<td>75</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Hostility</td>
<td>40</td>
<td>66</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Suspiciousness/persecution</td>
<td>3</td>
<td>37</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>Hallucinatory behavior</td>
<td>20</td>
<td>-1</td>
<td>70</td>
<td>-5</td>
</tr>
<tr>
<td>Unusual thought content</td>
<td>9</td>
<td>7</td>
<td>43</td>
<td>14</td>
</tr>
<tr>
<td>Conceptual disorganization</td>
<td>1</td>
<td>25</td>
<td>42</td>
<td>0</td>
</tr>
<tr>
<td>Depression</td>
<td>14</td>
<td>-3</td>
<td>5</td>
<td>66</td>
</tr>
<tr>
<td>Motor retardation</td>
<td>-9</td>
<td>6</td>
<td>-4</td>
<td>41</td>
</tr>
<tr>
<td>Somatic concerns</td>
<td>14</td>
<td>9</td>
<td>21</td>
<td>39</td>
</tr>
</tbody>
</table>

a The following BPRS items were excluded from the final factor solution because they were not found to load heavily on any factor when removed: grandiosity, blunted affect, emotional withdrawal, guilt feelings, mannerisms and posturing, and disorientation.

b Factor loadings are multiplied by 100 and rounded to the nearest integer. Significant loadings are highlighted in boldface type.

3. outcome=β0+β2 mediator + 
   β4 treatment + 
   β5 mediator × treatment

where β0, β2, and β4 are intercept terms; β1, β3, and β6 are coefficients associated with the magnitude of the treatment effect; β5 is a coefficient associated with the magnitude of the mediator effect; and β6 is a coefficient associated with the magnitude of the mediator/treatment interaction.

To test the hypothesis that change in psychosis is secondary to improvements in other aspects measured by the BPRS scale, we tested the hypothesis of whether the other factors were “mediators” of the change in core psychosis. In accordance with the equations, the factor representing core psychosis was treated as the “outcome” and others were treated as “mediators.” To see if the mediation was bidirectional, the other factors were treated as “outcome” and the psychosis factor was treated as a “mediator.” Mediation is said to occur when all of the following conditions are met:

1. The mediator is correlated with treatment.
2. The effect of treatment (β1) is significant in equation 1.
3. The effect of treatment (β2) is significant in equation 2.
4. The mediator has a significant effect (β3) in equation 3.

In cases where bidirectional mediation was found to occur, tests were performed to determine whether mediation was occurring more strongly in one direction than in the other. In these tests, the mediator and the outcome measure were both standardized, and then the coefficients associated with the mediator in equation 3 were compared across the two directions. A 95% confidence interval for each of these coefficients was constructed, and mediation was considered to occur equally in both directions if these two intervals overlapped. If they did not overlap, the dominant direction of the relationship could then be determined.

As another convergent test of the hypothesis, we addressed the issue of whether treatment has any effect on change in psychosis above and beyond reduction in nonspecific behavioral factors. A series of ANCOVAs (analysis of covariance) was performed to address this issue, in which measures of change in psychosis were used as the outcome variables, treatment was included as an explanatory variable, and change in nonspecific behavioral factors was included as a covariate. If the effect of treatment was found to be significant after adjustment for the relationship between the outcome measure and the covariate, then we could conclude that treatment had an additional effect on psychosis beyond merely reducing nonspecific behavioral factors.

Results

A total of 311 patients were included in the study and were randomly assigned to receive intramuscular olanzapine (N=131), intramuscular haloperidol (N=126), or intramuscular placebo (N=54). Baseline and 2-hour data were available for 286 (92%) patients, and data to assess changes at 24 hours were available for 273 (88%) patients, including 113 patients who received olanzapine, 112 who received haloperidol, and 48 who received placebo. There were no significant between-group baseline demographic or illness differences. The mean age of the patients was 38.2 years (SD=11.6, range=18–72); their mean age at onset of illness was 24.4 (SD=8.5, range=7–58). Thirty-nine percent (N=21) of the placebo group required the use of adjunctive benzodiazepines during the first 24 hours, compared to only 15% of the active treatment group (21 patients who received haloperidol and 16 who received olanzapine). The difference between the placebo and active treatment groups was significant (p<0.0001, Fisher’s exact test). The majority of patients who received adjunctive benzodiazepines received only one dose, and although more patients in the placebo group received benzodiazepines, there was no difference in the number and timing of benzodiazepine administration across the groups (χ²=1.02, df=2, p=0.60).

To determine the different dimensions of improvement (change in scores) after antipsychotic treatment, an unweighted least-squares approach was implemented, and three solutions, with three, four, and five factors, respectively, were obtained. Although all three solutions (Table 1) yielded similar findings, the four-factor solution was selected for further analysis for the following reasons: 1) all factors in this solution contained at least three significant item loadings, 2) this solution resulted in the least amount of overlap between the factors (i.e., the fewest cases where items loaded significantly on more than one factor), and 3) the items clustered within each factor could be interpreted in a clinically meaningful manner.
As Table 2 shows, factor analysis segregated items associated with agitation and excitement into the first factor (called agitation/excitement factor) and items associated with hostility into the second factor (hostility). The third factor contained items that measure psychosis (conceptual disorganization, hallucinatory behavior, and unusual thought content). The fourth factor brought in a few more distinct items (somatic concerns, depression, and motor retardation) that were not found to be significant in the three-factor solution. Relevant to our hypothesis, we found that change in agitation-related items segregated independent of improvement in psychosis-related items.

The analysis of factor scores revealed that both intramuscular olanzapine and intramuscular haloperidol show a significant effect on the psychosis factor at 24 hours, compared to placebo (Table 3).

In the case of intramuscular olanzapine, a significant effect on the psychosis factor was evident as early as 2 hours (t=2.80, df=284, p=0.01, with Bonferroni adjustment). Haloperidol was not differentiated from placebo at the 2-hour time point, although the difference approached significance (t=2.40, df=284, p=0.05, with Bonferroni adjustment). Intramuscular haloperidol and intramuscular olanzapine did not differ statistically in the degree of improvement on this factor at 2 hours or 24 hours.

The same question was also addressed by using the more conventional subscale that addresses core psychotic symptoms—the BPRS thought subscale. With that collection of items, intramuscular olanzapine was differentiated from placebo at 2 hours (t=2.84, df=283, p=0.01, with Bonferroni adjustment), and both intramuscular olanzapine and intramuscular haloperidol were differentiated from placebo at 24 hours (t=3.45, df=270, p=0.0007, with Bonferroni adjustment). To test if this improvement in the BPRS was secondary to improvement in nonspecific behavioral factors, we used covariance to adjust for the effects of improvement in nonspecific behavioral factors. There was a statistically significant effect of the active treatments on the BPRS thought subscale score after correction for the BPRS activation-hostility subscale score (overall model: F=14.54, df=3, 267, p<0.0001; independent effect of active treatment: t=−2.21, df=267 p=0.03). Similarly, there was a statistically significant effect of the active treatments on BPRS thought subscale scores after correction for Positive and Negative Syndrome Scale excitement component scores (overall model: F=11.79, df=3, 268, p<0.0001; independent effect of active treatment: t=2.58, df=268, p=0.01).

The mediation analysis revealed that the factors did explain the variance in each other (which suggests that they bear some relation); however, the mediation coefficients revealed bidirectionality. The size of the coefficient signifying mediation (β3 in equation 3) was as large when one considered the mediation of psychosis by changes in agitation-excitement (mean=0.36, SE=0.08) (p<0.0001, linear regression) as it was when one considered the mediation of agitation-excitement by psychosis (mean=0.34, SE=0.09) (p<0.0001, linear regression), thus refuting the hypothesis of unidirectional mediation of change in one by the other.

As a further confirmation of the overall hypothesis, the 2-hour change in agitation-excitement did not predict the 24-hour change in the psychosis factor (t=1.60, df=216, p=0.11). On the other hand, change in psychosis at the 2-hour mark predicted change in psychosis at the 24-hour mark (t=5.23, df=216, p<0.0001), suggesting that the early response in psychosis is distinct from changes in agitation-excitement and is continuous over time.

**Discussion**

There are no studies that have systematically examined the issue of early onset of antipsychotic response; however, results from a number of case series and uncontrolled studies have hinted at a rapid response to antipsychotic medications. Our study confirmed and extended these findings by testing them in a double-blind, placebo-controlled design and by showing that there are definite changes in measures of the psychotic component very early in treatment and that these changes are distinguishable from any changes in agitation or other nonspecific aspects of the acute presentation.

Our study has several limitations that must be taken into account in the interpretation and generalization of

---

**TABLE 3. Comparative Effects of Intramuscular Olanzapine, Haloperidol, and Placebo at 2 and 24 Hours After Treatment Onset in Patients With Acute Exacerbation of Psychosis**

<table>
<thead>
<tr>
<th>Score</th>
<th>Olanzapine Group Versus Placebo Group</th>
<th>Haloperidol Group Versus Placebo Group</th>
<th>Olanzapine Group Versus Haloperidol Group</th>
<th>Antipsychotics Groups Versus Placebo Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t</td>
<td>df</td>
<td>p</td>
<td>t</td>
</tr>
<tr>
<td>Psychosis factor score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-hour change</td>
<td>2.80</td>
<td>284</td>
<td>0.01</td>
<td>2.40</td>
</tr>
<tr>
<td>24-hour change</td>
<td>2.58</td>
<td>270</td>
<td>0.03</td>
<td>3.56</td>
</tr>
<tr>
<td>BPRS thought subscale score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-hour change</td>
<td>2.84</td>
<td>283</td>
<td>0.01</td>
<td>2.11</td>
</tr>
<tr>
<td>24-hour change</td>
<td>2.87</td>
<td>269</td>
<td>0.01</td>
<td>3.49</td>
</tr>
</tbody>
</table>

a Results of post hoc tests after application of a Bonferroni adjustment are reported (see Method section for details). Data were available at 2 hours for 113 patients in the olanzapine group, 112 patients in the haloperidol group, and 54 patients in the placebo group and at 24 hours for 113 patients in the olanzapine group, 112 patients in the haloperidol group, and 48 patients in the placebo group.
these results. First, the study was limited to patients who were demonstrating agitation or otherwise needed acute intervention. Although the results clearly demonstrate an improvement in psychosis, there still remains the question of whether quiet psychotic patients would also show a similar pattern of response. Agid et al. (4) showed that an unagitated psychotic group demonstrated a clear antipsychotic response by the first week, but they did not have the data to address responses earlier than the first week. This question remains an important, and tractable, issue for future clinical trials.

Second, the patients in this study had limited periods of washout, largely reflecting a clinical reality that required urgent action. In this situation, there is the possibility that previous treatment may have “primed” the patients for response, bringing them close to the threshold of response, and thus, the rather immediate response observed here may not be seen if the patients were drug naive or drug free for long periods. Although this explanation is possible, we believe it is unlikely, because almost all the previous studies (although not very well controlled) on this issue included drug-free or drug-naive patients and uniformly observed a rapid response (7, 9–16).

Third, the active treatment groups were less likely than the placebo group (15% versus 39%) to require the use of adjunctive benzodiazepines, as would be expected clinically. This difference works in favor of our hypothesis. That the antipsychotics were differentiated from placebo, despite twice as much use of adjunctive benzodiazepines (which provide sedation and behavioral control) in the placebo group further strengthens our claim of a specific and early antipsychotic action.

Fourth, the mode of administration of the antipsychotics in this study was intramuscular, which raises the question of whether a similar speed of onset of antipsychotic action would be observed had these medications been given orally. The intramuscular preparations certainly reach peak concentration earlier and have a higher level of bioavailability in the first few hours (21). For example, intramuscular olanzapine has a time to maximum plasma levels ($T_{\text{max}}$) of 0.5 hours, while oral olanzapine has a $T_{\text{max}}$ of 3.5 hours (22); intramuscular haloperidol has a $T_{\text{max}}$ of 40–60 minutes (23), while the $T_{\text{max}}$ for oral haloperidol is in the range of 4–6 hours (24). The difference is only a matter of a few hours after the first dose, and overall systemic exposure after a given dose is nearly the same (22, 25). Furthermore, a single oral dose of 15 mg of olanzapine showed 79%–80% dopamine D$_2$ receptor occupancy in the striatal and extrastriatal regions within 6 hours (26); a single oral dose of 4–7.5 mg of haloperidol showed greater than 80% occupancy of dopamine D$_2$ receptors within 3 hours (27). Thus, while drug availability after administration of intramuscular agents may be somewhat faster (21), we expect that a similar pattern is likely to be observed with oral administration of an appropriate dose.

Finally, the study examined only haloperidol, a typical antipsychotic, and olanzapine, an atypical antipsychotic, and therefore the extension of these findings to other agents is not warranted. However, if the meta-analysis of Agid et al. (4) is any guide, this phenomenon is likely to be relevant to other antipsychotics as well.

The study findings raise some interesting implications from a clinical as well as a theoretical perspective. Clinicians have always known about the acute effects of antipsychotic medications—that is why they are a staple in every emergency department. However, it is largely assumed that these acute effects are nonspecific or “behavioral” and that the antipsychotic effects do not occur until much later. The current data suggest a reconsideration of this viewpoint. Not only are these early effects apparent on the psychosis dimension, the magnitude of these early effects is not trivial. The degree of improvement reported in the first few days is actually rather striking. For example, in the study by Stern et al. (14), nearly 50% of the improvement seen at the end of 2 weeks was evident by the 3-day mark. The study by Glovinsky et al. (16) showed a significant change within the first 3 days of treatment and a relatively limited additional improvement in the weeks that followed. The meta-analysis by Agid et al. (4) showed that more improvement in psychosis was observed in the first week than in any week thereafter. The present study had an extension component wherein patients’ medication was switched to the respective oral preparation after the first day (28). It is interesting to note that although differentiation from baseline occurred with both active drugs within the first day, with olanzapine’s effects seen within the first 2 hours, there was no further significant improvement observed between days 1 and 5 of treatment in the olanzapine and haloperidol groups (28). Given this finding, we think that the early response is not just statistically significant but clinically relevant as well.

This study raises further questions about the hypothesis of delayed onset of antipsychotic action and leads us to propose an “early-onset hypothesis.” It is well recognized that substantial blockade of the dopamine system by antipsychotic medications happens within the first few hours of treatment (26, 27). Because it was assumed all along that clinical response was delayed in onset, this dissociation between early occupancy and delayed onset of antipsychotic action was seen as a paradox. To explain this paradox, several explanations have been proposed. A widely considered explanation was that of a depolarization blockade. According to that hypothesis, the acute blockade of dopamine receptors led to an increased firing of dopamine neurons, which finally resulted in an exhaustion and depolarization blockade of the neurons 2–3 weeks down the line (29–31). This terminal event was seen as the mediator responsible for the delayed onset of antipsychotic response (31). Along slightly different lines, it has been proposed that the dopamine blockade leads to changes in gene expression, which over the ensuing days...
and weeks, lead to synaptic plastic changes, which are then seen as the proximal mediators of the onset of antipsychotic response (32). Although the depolarization blockade and synaptic plastic changes remain important empirical phenomena and may have a role in the longer-term consequences of antipsychotics, in light of the fact that the onset of antipsychotic response precedes these changes, their role as critical mediators of antipsychotic response may need to be reconsidered.

The realization of the early onset of antipsychotic response calls for a different kind of explanation than that suggested by the mechanisms proposed previously. A study by Abi-Dargham et al. (33) may provide an interesting clue in this regard. Using alpha-methyl-para-tyrosine, an agent that causes acute dopamine depletion (a fact verified by brain imaging in that study), they showed a robust antipsychotic response within the first 48 hours. Thus, the interruption of dopamine transmission at D2 receptors, whether by blockade or depletion, may be the immediate mediator of antipsychotic response (1). If that is the case, theories that relate alterations in dopamine transmission to their proximate effects on reward, motivational salience, and prediction error (34–38) may have more bearing on antipsychotic effect than accounts that focus on mechanisms that require 2–3 weeks for onset.

In conclusion, the study demonstrated that in acutely agitated psychotic patients with a schizophrenia spectrum disorder, the acute administration of a typical or atypical antipsychotic leads to a robust and independent improvement in psychotic symptoms. This improvement is observable as early as 2 hours after treatment with olanzapine and definitely by 24 hours with both olanzapine and haloperidol. This improvement is distinct from a placebo response and is not secondary to drug-induced changes in anxiety, agitation, or other nonspecific factors. These findings question the well-accepted “delayed-onset” hypothesis of antipsychotic action and call for a reconsideration of some of the clinical practices and scientific theories that have come in its wake.

Received Jan. 13, 2004; revision received April 9, 2004; accepted May 10, 2004. From the Centre for Addiction and Mental Health; the Department of Psychiatry, University of Toronto, Toronto; Eli Lilly and Co., Indianapolis; and Eli Lilly Canada, Toronto. Address correspondence and reprint requests to Dr. Kapur, Centre for Addiction and Mental Health, 250 College St., Toronto, Ont., Canada MSR 1T8; shitij.kapur@camh.net (e-mail).

The data used in this analysis were provided by Eli Lilly and Co. No financial support was sought or received from Eli Lilly for this work. Dr. Lindborg and Dr. Jones are employees of Eli Lilly. Dr. Kapur’s research in this area is supported by a Canada Research Chair.

References

1. Kapur S, Remington G: Dopamine D2 receptors and their role in atypical antipsychotic action: still necessary and may even be sufficient. Biol Psychiatry 2001; 50:873–883
33. Abi-Dargham A, Rodenhiser J, Printz D, Zea-Ponce Y, Gil R, Kegeles LS, Weiss R, Cooper TB, Mann JJ, Van Heertum RL, Gorman JM, Laruelle M: Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. Proc Natl Acad Sci USA 2000; 97:8104–8109
Risperidone and Haloperidol in First-Episode Psychosis: A Long-Term Randomized Trial

Nina Schooler, Ph.D.
Jonathan Rabinowitz, Ph.D.
Michael Davidson, M.D.
Robin Emsley, M.D.
Philip D. Harvey, Ph.D.
Lili Kopala, M.D.
Patrick D. McGorry, M.D.
Ilse Van Hove, M.Sc.
Marielle Eerdekens, M.D.
Wim Swyzen, M.D.
Goedele De Smedt, M.D.
Early Psychosis Global Working Group

Objective: The first episode of psychotic illness is a key intervention point. The initial experience with medication can affect willingness to accept treatment. Further, relapse prevention is a treatment cornerstone during the first years of illness because active psychotic illness may affect lifetime outcomes. Thus, initial treatment of active symptoms and subsequent relapse prevention are central goals of pharmacotherapy. This study compared long-term effectiveness of risperidone versus haloperidol in first-episode psychosis patients.

Method: First-episode psychosis patients (N=555, mean age=25.4 years) participated in a double-blind, randomized, controlled flexible-dose trial that compared risperidone (mean modal dose=3.3 mg) and haloperidol (mean modal dose=2.9 mg). The median treatment length was 206 days (maximum=1,514).

Results: Positive and Negative Syndrome Scale scores and Clinical Global Impressions ratings improved significantly relative to baseline, with no significant differences between groups. Three-quarters of the patients achieved initial clinical improvement, defined as >20% reduction in total Positive and Negative Syndrome Scale score. However, among those who achieved clinical improvement, 42% of the risperidone group experienced a relapse compared with 55% of the haloperidol group. The median time to relapse was 466 days for risperidone-treated subjects and 205 days for those given haloperidol. These differences were statistically significant based on Kaplan-Meier survival analysis. Adverse effects distinguished the treatments: there were significantly more extrapyramidal signs and symptoms and adjunctive medication use in the haloperidol group and greater prolactin elevation in the risperidone group. There was less weight gain with haloperidol initially but no significant differences between groups at endpoint.

Conclusions: Relatively low doses of antipsychotic drugs lead to significant symptom amelioration in the majority of first-episode psychosis patients. In the long term, risperidone prevents relapse in more patients and for a longer time and also induces less abnormal movements than haloperidol.

Following the first episode of psychosis, treatment with antipsychotic medication is associated with rapid improvement of symptoms in a majority of individuals (1–7). First-episode patients appear to respond to relatively low doses of antipsychotic medication (4, 8) and manifest high sensitivity to extrapyramidal signs and symptoms (9). Unfortunately, because of the chronic undulating course of psychosis in schizophrenia, the majority of first-episode patients experience a relapse within the first year after clinical improvement or remission, either because of medication discontinuation or despite continuous treatment (10). Although treating acute symptoms and preventing relapse are important at any time during the illness and at any age, it is particularly critical in adolescents and young adults and during the first few years of the illness. This is because the illness may be more active in the initial phases, with frequent and distinct cycles of remission and exacerbations (4). Also, late adolescents and early adulthood are critical years for social and vocational development. Hence, illness control might have an impact on life-long outcomes. Reducing the number of relapses and increasing the time spent with few or no symptoms is therefore a major goal of pharmacological treatment.

A number of meta-analyses have tried to determine to what extent the novel antipsychotics are superior to the older-generation antipsychotics (11, 12), specifically in terms of maintaining symptomatic improvement (13) and preventing relapse (14). A previous trial designed to compare the efficacy of risperidone versus haloperidol in preventing relapse in stable outpatients with chronic schizophrenia or schizoaffective disorder (15) reported significantly fewer occurrences of relapse with risperidone and after a significantly longer time of treatment. Patients treated with risperidone were about half as likely to experience relapse than those treated with haloperidol. The current study compared the efficacy of risperidone and haloperidol in preventing relapse in first-episode psychosis patients. We also compared long-term symptom efficacy and adverse effect profiles of the medications.
TABLE 1. Demographic and Clinical Characteristics of Patients With First-Episode Psychosis Randomly Assigned to Double-Blind Treatment With Risperidone or Haloperidol

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Treatment Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Risperidone</td>
</tr>
<tr>
<td></td>
<td>(N=278)</td>
</tr>
<tr>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Male</td>
<td>196</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>205</td>
</tr>
<tr>
<td>Black</td>
<td>36</td>
</tr>
<tr>
<td>Hispanic</td>
<td>10</td>
</tr>
<tr>
<td>Other</td>
<td>27</td>
</tr>
<tr>
<td>DSM-IV diagnosis</td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>152</td>
</tr>
<tr>
<td>Schizoaffective disorder</td>
<td>17</td>
</tr>
<tr>
<td>Schizophreniform disorder</td>
<td>109</td>
</tr>
<tr>
<td>No previous antipsychotic exposure</td>
<td>94</td>
</tr>
<tr>
<td>Age (years)</td>
<td>25.2</td>
</tr>
<tr>
<td>Age at onset of first psychotic symptoms (years)</td>
<td>20.89</td>
</tr>
<tr>
<td>Male subjects</td>
<td>22.89</td>
</tr>
<tr>
<td>Female subjects</td>
<td></td>
</tr>
</tbody>
</table>

Method

The study, sponsored by Johnson & Johnson Pharmaceutical Research and Development, enrolled patients in 11 countries between November 1996 and January 2000 with planned treatment until the last enrolled participant completed 2 years of treatment. The study was conducted in accordance with good clinical practice after it was approved by the local institutional review boards.

Subjects

Consenting 16–45-year-old patients were enrolled into the trial if they 1) met Structured Clinical Interview for DSM-IV criteria for schizophrenia, schizophreniform disorder, or schizoaffective disorder for no more than 1 year during which period they had no more than two psychiatric hospitalizations for psychosis; and 2) had less than 12 weeks of cumulative exposure to antipsychotics and required antipsychotic treatment upon enrollment into the trial. Patients were excluded from the trial for any of the following reasons: 1) meeting DSM-IV criteria for another axis I diagnosis, including substance dependence or abuse; 2) needing another antipsychotic medication at enrollment; or 3) having a serious or unstable medical illness.

Study Design and Procedures

Before entering the study, subjects provided informed consent after the procedure had been fully explained. Subjects admitted to this double-blind trial were randomly allocated to receive either risperidone or haloperidol according to a 1:1 randomization scheme balanced by site. Before administration of trial medication patients had a 3- to 7-day drug washout period that was waived for extremely ill patients. Subjects in both treatment groups started with a once daily dose of 1 mg that could be increased to 2 mg/day on day 4 and thereafter by 1 mg/day each week, up to a maximum daily dose of 4 mg. In exceptional cases (i.e., subjects showing insufficient response in whom not more than mild extrapyramidal signs and symptoms were observed at 4 mg/day), the dose could then be increased further by 1 mg a week up to a maximum daily dose of 8 mg. Concomitant psychotropic medications allowed were those addressing extrapyramidal signs and symptoms; chloral hydrate, zolpidem, or flurazepam for sleep; and lorazepam for agitation.

Assessments

Outcomes were measured in five domains: 1) relapse, 2) psychopathology, 3) safety, 4) quality of life, and 5) neurocognitive functioning (the latter two are reported in separate manuscripts). Psychopathology was assessed using the Positive and Negative Syndrome Scale (16) and Clinical Global Impression (CGI) severity and change scales (17).

Relapse was examined among patients who reached clinical improvement (decrease of more than 20% on total Positive and Negative Syndrome Scale score) and was defined according to Csernansky et al. criteria (15) as any one of the following occurring after clinical improvement: 1) 25% or more increase in score on the Positive and Negative Syndrome Scale (or a 10-point increase if initial score was 40 or less); 2) CGI change rating of "much worse" or "very much worse"; 3) deliberate self-injury (as a reported adverse event); 4) emergence of clinically significant suicidal or homicidal ideation (as a reported adverse event); or 5) violent behavior resulting in significant injury to another person or significant property damage (as a reported adverse event).

Abnormal involuntary movements were assessed with the Extrapyramidal Symptom Rating Scale (18). Adverse effects were recorded with standard recording forms.

Assessments with the Positive and Negative Syndrome Scale, CGI, and Extrapyramidal Symptom Rating Scale were completed weekly during the first 4 weeks of the trial and then every 4 weeks for the next 5 months. During months 6–15, the instruments were completed every 2 months and every 3 months thereafter. Follow-up evaluations were conducted until the last patient enrolled had completed 2 years of treatment. The blind was broken when the study ended.

Statistical Analysis

Five hundred fifty-nine patients from 11 countries were randomly assigned to receive either haloperidol or risperidone. Three patients assigned to risperidone and one patient assigned to haloperidol did not receive study medication and were thus excluded from the analysis. Therefore, 278 patients treated with risperidone and 277 with haloperidol were included in the analysis. The data from all randomized and treated patients were analyzed for safety. Before breaking the blind, 11 subjects receiving risperidone and 10 subjects receiving haloperidol, all from the same site, were excluded from the efficacy analyses because of violations of good clinical practice. Exclusion of these data did not change the findings. Baseline characteristics and duration of treatment were compared between the two groups by analysis of variance or chi-square test for categorical variables and summarized by descriptive statistics.

Differences between the groups in the degree of change from baseline in scores on the Positive and Negative Syndrome Scale and Extrapyramidal Symptom Rating Scale were evaluated with analysis of covariance after we controlled for baseline scores and tested for center-by-treatment interactions. For the Extrapyramidal Symptom Rating Scale, change for each patient was examined from baseline to maximum score at any time point. Differences on CGI change scale were also tested but with no baseline controls.

Further analysis of dyskinesia used the Extrapyramidal Symptom Rating Scale dykinetic movement scale to operationalize dyskinesia on the basis of criteria of Schooler and Kane (19). Those criteria were originally developed for the Abnormal Involuntary Movement Scale (20), which contains almost identical items on a 5-point scale. Emergent dyskinesia was defined according to the Schooler and Kane criteria as an increase from baseline of 3 points or more on one item or 2 points or more on
two items of the seven-item dyskinetic movement scale. Persistent dyskinesia was defined as emergent dyskinesia that met the criteria on two or more consecutive visits.

Relapse rates, which were part of the planned analysis, were analyzed according to the aforementioned Csernansky et al. criteria (15). As planned, the relapse analysis included only those patients who had reached clinical response as defined as a >20% decrease in score on the total Positive and Negative Syndrome Scale. A dichotomous variable was created that assigned a value of 1 once a patient experienced a relapse. Time to relapse was calculated as the number of days elapsing from clinical improvement to the first relapse using Kaplan-Meier survival analysis. Cox regression analysis was performed to test for possible interaction effects of treatment group and study center. The differences in time to relapse between treatment groups were analyzed by using Cox proportional hazards model and log rank test after we controlled for center. Follow-up evaluations for endpoint analyses ceased after the discontinuation of treatment. The analysis of time to relapse was therefore censored at the time of treatment discontinuation if it occurred before relapse. All statistical tests were two-tailed. All analyses tested for study center-by-treatment interactions; no such interactions were found.

Results

The characteristics of the 555 patients in the two treatment groups were similar (Table 1). About half had a diagnosis of schizophrenia, 70% were male, and their mean age was 25 years.

Subjects were treated with trial medication for a median of 192 days (range=2–1,502) in the risperidone group and for a median of 218 days (range=1–1,514) in the haloperidol group (Mann-Whitney z=0.116, p=0.90). The mean modal total daily dose was 3.3 mg for risperidone and 2.9 mg for haloperidol. The most commonly taken daily dose for each of the drugs (mode dose) was 3 mg. Two hundred eighteen subjects (117 in the risperidone group and 101 in the haloperidol group) discontinued double-blind treatment prematurely. As can be seen in Table 2, there were no significant differences between groups in overall discontinuation or specific reasons for discontinuation.

After 3 months, 73.6% (N=192) of patients randomly assigned to risperidone showed clinical improvement (>20% decrease in Positive and Negative Syndrome Scale score) as did 76.2% (N=199) of those receiving haloperidol ($\chi^2$=0.50, df=1, p=0.48). Kaplan-Meier survival analysis found that at study endpoint slightly more than three-quarters of the patients in each group (risperidone, N=197 [75.5%]; haloperidol, N=203 [77.8%]) met the predefined clinical improvement criterion, with a median time to clinical improvement of 26 days in the risperidone group and 22 days in the haloperidol group (log rank=1.49, p=0.22). As shown in Table 3, both groups showed clinical improvement according to Positive and Negative Syndrome Scale scores and CGI ratings, with no significant differences between the risperidone and haloperidol groups. According to CGI change ratings, 50.8% of the patients were “much” or “very much improved,” and 81.4% were at least minimally improved.

Among those patients who achieved clinical improvement (risperidone, N=197; haloperidol, N=203), there were significantly fewer relapses in the risperidone group (42.1%) than in the haloperidol group (54.7%). The time to relapse for the risperidone group was significantly longer than for the haloperidol group (risperidone median=466 days; haloperidol median=205 days). Figure 1 presents Kaplan-Meier plot of time from clinical response until relapse, the illustrated difference between the curves is highly significant (log rank=7.10, df=1, p=0.008). Significant differences between the groups emerged by 145 days, at which time there were 58 events of relapse in the risperidone group and 80 in the haloperidol group (mean days to relapse=114 for risperidone, 102 for haloperidol) (log-rank test comparing survival curves, p<0.04).

Safety

Treatment-emergent extrapyramidal signs and symptoms were significantly more frequent and more severe in the haloperidol-treated group as reflected by the scores on the Extrapyramidal Symptom Rating Scale (Table 4). There was significantly less emergent dyskinesia in the risperidone group than in the haloperidol group but no significant difference in persistent dyskinesia. On the specific Extrapyramidal Symptom Rating Scale subscales, the risperidone group had significantly lower maximum change in score from baseline on total, parkinsonism, and parkinsonism dystonia (due primarily to the difference in parkinsonism) symptoms. Significantly lower akathisia scores were seen in the risperidone-treated group as well as a tendency for lower dyskinesia scores.

---

### TABLE 2. Discontinuation Rates and Reasons Among Patients With First-Episode Psychosis Randomly Assigned to Double-Blind Treatment With Risperidone or Haloperidol

<table>
<thead>
<tr>
<th>Discontinuation Variable</th>
<th>Risperidone (N=278)</th>
<th>Haloperidol (N=277)</th>
<th>Total (N=555)</th>
<th>Analysisa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total discontinuing treatment</td>
<td>117</td>
<td>42.1</td>
<td>101</td>
<td>36.5</td>
</tr>
<tr>
<td>Reasons for discontinuation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse event</td>
<td>15</td>
<td>5.4</td>
<td>17</td>
<td>6.1</td>
</tr>
<tr>
<td>Insufficient response</td>
<td>25</td>
<td>9.0</td>
<td>16</td>
<td>5.8</td>
</tr>
<tr>
<td>Subject ineligible to continue the trial</td>
<td>1</td>
<td>0.4</td>
<td>3</td>
<td>1.1</td>
</tr>
<tr>
<td>Subject lost to follow-up</td>
<td>20</td>
<td>7.2</td>
<td>16</td>
<td>5.8</td>
</tr>
<tr>
<td>Subject noncompliant</td>
<td>5</td>
<td>1.8</td>
<td>8</td>
<td>2.9</td>
</tr>
<tr>
<td>Subject withdrew consent</td>
<td>34</td>
<td>12.2</td>
<td>32</td>
<td>11.6</td>
</tr>
<tr>
<td>Other</td>
<td>17</td>
<td>6.1</td>
<td>9</td>
<td>3.2</td>
</tr>
</tbody>
</table>

a For risperidone versus haloperidol.
Similarly, there were also differences in the rates of patients who received concomitant medications as treatments for extrapyramidal signs and symptoms. For the haloperidol- and risperidine-treated patients, 49.5% (N=137) and 41.7% (N=116), respectively, received anticholinergic agents (χ²=3.34, df=1, p<0.07); 61.7% (N=171) and 54.7% (N=152) received benzodiazepines to control agitation as well as extrapyramidal signs and symptoms (χ²=2.84, df=1, p<0.10); and 10.5% (N=29) and 5.0% (N=14) received beta blocking agents to control akathisia (χ²=5.73, df=1, p<0.02). Forty-six patients expressed suicidal ideations during the trial. This was 9.4% (N=26) of the haloperidol group, in which there were three completed suicides, versus 7.2% (N=20) of the risperidone group, in which no suicides were completed.

Significantly more weight gain was observed in the risperidone group early in treatment. At the third month of treatment the risperidone group (N=180) had gained an average 4.6 kg (SD=4.96) and the haloperidol group (N=204) 3.5 kg (SD=4.42) (t=5.41, df=358, p=0.03). At endpoint the difference in weight gain between the treatment groups was no longer significant (risperidone [N=211]: mean=7.5 kg, SD=9.29; haloperidol [N=204]: mean=6.5 kg, SD=8.86) (t=1.13, df=413, p=0.26).

There were no notable differences between the treatment groups for vital signs, reported adverse events, or ECG parameters. On laboratory parameters, the only notable difference was that maximum prolactin levels (ng/ml) were higher in the risperidone group (women [N=73]: mean=73.69, SD=53.18; men [N=185]: mean=34.08, SD=21.90) than in the haloperidol group (women [N=71]: mean=48.16, SD=47.82; men [N=178]: mean=21.81, SD=14.54) (for women: t=3.03, df=142, p<0.003; for men: t=6.31, df=361, p=0.0001). There were abnormal prolactin values (males >18 ng/ml; females >25 ng/ml) in 73.8% (N=189 of 256) of the risperidone patients and in 49.8% (N=124 of 249) of the haloperidol-treated patients. Prolactin-related adverse effects were reported in 14 risperidone-treated patients and one haloperidol-treated patient. All patients with prolactin-related adverse effects had abnormal prolactin levels. Specifically, among the risperidone patients, there were three patients with gynecomastia, six with hyperprolactinemia, and six with galactorrhea. One patient had both hyperprolactinemia and galactorrhea. There was one case of hyperprolactinemia in the haloperidol group. Moderate hyperglycemia was reported as an adverse effect in one risperidone-treated subject.

Discussion

This study both confirms findings regarding treatment of the first episode of schizophrenia and extends our understanding of the role of medication during this critical period of the illness. In the present study we found initial symptom improvement in a carefully defined first-episode patient group treated with low doses of either a conventional antipsychotic (haloperidol) or an atypical medication (risperidone). This finding is in agreement with results of studies with older antipsychotic medications (1, 8, 21). It is also in accord with results of studies of second-generation antipsychotic medications (5, 7). However, even when dosage is appropriately managed for first-episode patients, haloperidol (at a mean dose of 2.9 mg/day) is associated with significantly greater acute extrapyramidal signs and symptoms and a greater need for concomitant medications to treat those side effects than risperidone. The favorable response rate to antipsychotic medication in this study was similar to rates reported by other trials treating recent-onset psychosis (7, 21).

The further and unique contributions of this study are a function of the long duration of treatment and the fact that it was a randomized, double-blind trial. To our knowledge this is the longest such trial to compare an older an-

---

**TABLE 3. Psychopathology at Baseline and Endpoint Among Patients With First-Episode Psychosis Randomly Assigned to Double-Blind Treatment With Risperidone or Haloperidol**

<table>
<thead>
<tr>
<th>Psychopathology Measure</th>
<th>Risperidone</th>
<th>Haloperidol</th>
<th>Analysisa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Mean SE 95% CI</td>
<td>N Mean SE 95% CI</td>
<td>F (df=1, 516) p</td>
</tr>
<tr>
<td><strong>Positive and Negative Syndrome Scale</strong> Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>266 83.7 1.24 –23.9 to –21.0</td>
<td>267 81.1 1.23 –23.4 to –17.8</td>
<td>0.47 0.49</td>
</tr>
<tr>
<td>Change at endpoint</td>
<td>264 –21.0 1.46 –23.9 to –21.0</td>
<td>264 –20.6 1.43 –23.4 to –17.8</td>
<td></td>
</tr>
<tr>
<td><strong>Positive symptoms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>266 20.5 0.37 –7.4 to –5.8</td>
<td>267 20.0 0.39 –7.9 to –6.1</td>
<td>2.30 0.13</td>
</tr>
<tr>
<td>Change at endpoint</td>
<td>264 –6.6 0.43 –7.4 to –5.8</td>
<td>264 –7.0 0.48 –7.9 to –6.1</td>
<td></td>
</tr>
<tr>
<td><strong>Negative symptoms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>266 22.0 0.44 –5.7 to –3.9</td>
<td>267 21.0 0.43 –5.1 to –3.3</td>
<td>0.00 0.98</td>
</tr>
<tr>
<td>Change at endpoint</td>
<td>264 –4.8 0.44 –5.7 to –3.9</td>
<td>264 –4.2 0.44 –5.1 to –3.3</td>
<td></td>
</tr>
<tr>
<td><strong>General psychopathology</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>266 41.3 0.66 –11.1 to –8.1</td>
<td>267 40.1 0.63 –10.8 to –7.8</td>
<td>0.25 0.62</td>
</tr>
<tr>
<td>Change at endpoint</td>
<td>264 –9.6 0.77 –11.1 to –8.1</td>
<td>264 –9.3 0.75 –10.8 to –7.8</td>
<td></td>
</tr>
<tr>
<td><strong>Clinical Global Impression</strong> change score</td>
<td>263 2.69 0.07 2.55 to 2.83</td>
<td>264 2.62 0.07 2.48 to 2.76</td>
<td>0.56 0.45</td>
</tr>
</tbody>
</table>

*Comparison of treatment groups in terms of change from baseline.*
tipsychotic medication and an atypical agent. As a result, we were able to study the process of relapse after the initial response. The finding of an increased risk of relapse in haloperidol-treated subjects relative to those treated with risperidone is of great clinical significance. Because treatment for all participants continued until the last participant enrolled had an opportunity for 2 years of treatment, the earliest enrolled patients could be treated and followed for almost 6 years. Using this design we observed a significant increase in time to relapse for risperidone that emerged as early as 100 days after first clinical improvement and persisted until the end of the trial. The substantial delay in relapse with risperidone would not have been detected in a brief trial. Inspection of the survival curves also reveals that the magnitude of the difference does not decrease over time. Despite this substantial clinical advantage for risperidone, the extended observation period makes it clear that relapse does occur but that risperidone serves to delay and prevent that event.

The most frequent and disturbing adverse effects associated with treatment—specifically extrapyramidal signs and symptoms, including emergent dyskinesia and akathisia—were less prevalent in the risperidone-treated patients than in the haloperidol-treated patients. The lower prevalence of extrapyramidal signs and symptoms in the risperidone-treated patients has possible implications regarding adherence to medications, since extrapyramidal signs and symptoms are associated with poorer compliance (22). There are also possible implications for suicide prevention, since suicidality may also be related to extrapyramidal signs and symptoms (23). Reports of suicidal ideation as an adverse event did not differ significantly between the two treatment groups. Three completed suicides all occurred in the haloperidol group; this represents 1.08% over the course of the trial. The expected rate based on epidemiological studies of first-episode schizophrenia is estimated to be between 1% and 2% for the first year after initial hospitalization (24).

The novel atypical antipsychotics produce less extrapyramidal signs and symptoms than typical ones, but as seen in the present study, they are associated with a greater risk of weight gain (25). We found weight gain with both haloperidol and risperidone: significantly more gain with risperidone at the 3-month point (the first time point at which weight was measured following baseline), but no significant difference at endpoint. Furthermore, it appears that younger, recent-onset psychosis patients are more vulnerable to antipsychotic-induced weight gain than older patients with more chronic illness. Since most recent-onset patients are targeted with novel atypical agents, weight gain is a major consideration in selecting the specific atypical drug. A recent conference on antipsychotic drugs, obesity, and diabetes (26) concluded that clozapine and olanzapine are associated with the greatest
TREATMENT OF FIRST-EPISODE PSYCHOSIS

TABLE 4. Extrapyramidal Symptoms in Patients With First-Episode Psychosis Randomly Assigned to Double-Blind Treatment With Risperidone or Haloperidol

<table>
<thead>
<tr>
<th>Symptom Measure</th>
<th>Risperidone</th>
<th>Haloperidol</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N With Symptom</td>
<td>%</td>
<td>N With Symptom</td>
</tr>
<tr>
<td>Dyskinesiaa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>278</td>
<td>3</td>
<td>276</td>
</tr>
<tr>
<td>Emergent</td>
<td>278</td>
<td>23</td>
<td>276</td>
</tr>
<tr>
<td>Persistent</td>
<td>278</td>
<td>5</td>
<td>276</td>
</tr>
<tr>
<td>Extrapyramidal symptoms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>275</td>
<td>3.72</td>
<td>5.09</td>
</tr>
<tr>
<td>Parkinsonism, dystonia</td>
<td>275</td>
<td>3.28</td>
<td>4.58</td>
</tr>
<tr>
<td>Dystonia</td>
<td>275</td>
<td>0.34</td>
<td>0.98</td>
</tr>
<tr>
<td>Parkinsonism</td>
<td>275</td>
<td>3.12</td>
<td>4.45</td>
</tr>
<tr>
<td>Dyskinesia</td>
<td>275</td>
<td>0.82</td>
<td>2.00</td>
</tr>
<tr>
<td>Akathisia (single item)</td>
<td>275</td>
<td>0.61</td>
<td>1.17</td>
</tr>
</tbody>
</table>

a Schoeler and Kane criteria (19).

b Scores represent maximum change on Extrapyramidal Symptom Rating Scale score from baseline.

weight gain and that risperidone and quetiapine have intermediate effects.

Risperidone treatment was associated with significantly greater elevation of prolactin in both men and women compared with haloperidol. Seventy-four percent of risperidone patients developed abnormal prolactin levels at some point during the trial, compared with 49% of the haloperidol-treated patients. These elevated prolactin levels were rarely associated with reports of prolactin-related adverse events.

A limitation of randomized controlled trials is that individuals who agree to participate in such trials may not be representative of the patients treated in routine clinical practice or of the general population of individuals suffering from schizophrenia, and therefore the results are not generalizable. To address this limitation, we compared the baseline characteristics of the subjects enrolled in this clinical trial to a large-scale epidemiological sample of first-episode patients previously collected in the United States (the Suffolk County Mental Health Project) (27). We found that 33% (N=59) of the epidemiological sample would not have met inclusion criteria for the present drug trial (because of antidepressant treatment, N=26; current substance abuse, N=18; recent suicide attempt, N=9; or for more than one reason, N=6). There were no significant differences between the two study groups on age at onset, age, gender, or premorbid functioning. Drug trial patients had more severe clinical symptoms, slightly lower CGI ratings, and less formal education than those in the epidemiological study. Although the epidemiological sample itself was from a limited geographic area, the general characteristics of these two study groups were similar on several key variables, supporting the generalizability of the findings. A second threat to generalizability concerns overall retention in the trial. As shown in Table 2 there were no treatment differences in withdrawal from the trial because of competing risks. Discontinuation due to relapse represented an outcome of the study. Overall, the percent discontinued was just under 40%.

The study raises many questions that call for further research. The first has to do with the determinants of relapse. In this trial, medication adherence was closely monitored. However, with oral medication in outpatients, lack of adherence to medication may have occurred. Robinson et al. (3), who followed first-episode patients regardless of treatment status, found that nonadherence was the strongest predictor of relapse in their study. It is clear that oral risperidone delays relapse substantially. Whether control for adherence through use of a long-term injectable form of medication, which is now available for risperidone, would delay relapse further needs to be investigated. A second important question is whether the delay of relapse seen in two long-term studies with risperidone is specific to risperidone or is a class effect that will be seen with other atypical antipsychotics such as olanzapine, quetiapine, ziprasidone, and aripiprazole. However, the combination of greater delay of relapse and a reduced emergent dyskinesia strongly support the use of risperidone early in a schizophrenic illness.

In summary, this report of the largest treatment study of first-episode psychosis patients, which also featured an unusually long duration, contributes to the emerging body of knowledge on the treatment response of first-episode psychosis patients and the comparative efficacy and safety of the atypical and conventional antipsychotic drugs. The present study results demonstrate a clear advantage for risperidone relative to haloperidol in terms of relapse prevention and in diminished incidence of extrapyramidal signs and symptoms.

Acknowledgments

The Early Psychosis Global Working Group comprised the following investigators: Australia—P. McGorry (Melbourne); T. Lambert (Bentley); J. Kulkarni (Dandenong); Austria—W. Fleischhacker (Innsbruck); Canada—D. Addington (Calgary); L. Kopala (Halifax, N.S.); R. Williams (Calgary); G. Chouinard (Montreal); A. Labelle

http://ajp.psychiatryonline.org

Am J Psychiatry 162:5, May 2005
References


A Double-Blind, Placebo-Controlled Trial of Sibutramine for Olanzapine-Associated Weight Gain

David C. Henderson, M.D.
Paul M. Copeland, M.D.
Tara B. Daley, M.P.H.
Christina P. Borba, M.P.H.
Corrine Cather, Ph.D.
Dana D. Nguyen, Ph.D.
Pearl M. Louie, M.D.
A. Eden Evins, M.D.
Oliver Freudenreich, M.D.
Doug Hayden, M.A.
Donald C. Goff, M.D.

Objective: Weight gain is commonly observed with olanzapine treatment and can increase the risk for obesity, cardiovascular disease, hypertension, and diabetes mellitus. This study examined the effectiveness of sibutramine, an approved weight loss agent, in overweight and obese subjects taking olanzapine for schizophrenia or schizoaffective disorder.

Method: Each subject had a DSM-IV diagnosis of schizophrenia or schizoaffective disorder, had been taking a stable dose of olanzapine for at least 4 months, and had a body mass index of ≥30 kg/m² or ≥27 kg/m² plus at least one cardiovascular risk factor. In a 12-week double-blind, randomized, placebo-controlled study, 37 subjects received placebo or sibutramine (up to 15 mg/day). For the first 8 weeks all subjects participated in weekly group sessions focused on nutrition and behavioral modification.

Results: The sibutramine and placebo groups had no significant baseline differences on age, gender, education, ethnicity, diagnosis, weight, body mass index, and blood pressure. At week 12 the sibutramine group had significantly greater losses than the placebo group in weight (mean=8.3 lb, SD=2.4, versus mean=1.8 lb, SD=1.6), waist circumference, body mass index, and hemoglobin A1c. There were no significant differences on most side effects, although the sibutramine group exhibited a mean increase in systolic blood pressure of 2.1 mm Hg (SD=8.5), and anticholinergic side effects and sleep disturbances were at least twice as common in the sibutramine group.

Conclusions: Sibutramine was an effective and well-tolerated adjunct to behavior modification for weight loss in patients with schizophrenia and schizoaffective disorder being treated with olanzapine.
The adverse effects of antipsychotic-induced weight gain may be greatest in men, the elderly, patients with higher body mass indexes at baseline, or patients who experience greater degrees of weight gain (11). Thus, the benefits of atypical antipsychotics, including a decreased rate of mortality from suicide, may be offset by the consequences of antipsychotic-induced weight gain (11). Aside from medical complications, weight gain is also an emotionally distressing side effect that contributes to nonadherence with antipsychotic treatment (12–14). Allison et al. (12) reported that patients treated with antipsychotic medications who gain weight have a reduced quality of life, poorer self-reported general health, and decreased vitality.

While switching to a more weight-neutral atypical antipsychotic agent offers promise in halting or reversing weight gain associated with an antipsychotic agent, many patients and their clinicians are reluctant to risk a worsening or return of psychotic symptoms. As a result, various agents have been proposed as adjunctive treatments to attenuate antipsychotic-induced weight gain. In a 6-week double-blind, placebo-controlled study of 26 inpatients, the addition of reboxetine, a selective norepinephrine reuptake inhibitor, to olanzapine at a dose of 10 mg/day resulted in significantly less weight gain (mean=2.5 kg, SD=2.7, versus mean=5.5 kg, SD=3.1) and significant improvement in the Hamilton Depression Rating Scale score (15). In another study, coadministration of fluoxetine was ineffective in diminishing olanzapine-induced weight gain in first-episode schizophrenia patients (16).

Sibutramine hydrochloride, a weight loss agent affecting both serotonin and norepinephrine reuptake, was introduced into the U.S. market in 1997 (17). The hypophagic effect of sibutramine is thought to be mediated, in part, through activation of the serotonin 5-HT2C receptor (18). Sibutramine has been shown to be an effective and well-tolerated weight loss agent for obesity (19). In a 44-week randomized, placebo-controlled trial, sibutramine produced weight loss and a decrease in waist circumference. Fasting triglyceride levels decreased and levels of high-density lipoprotein (HDL) cholesterol increased; these changes are often associated with weight loss (20). In a 1-year placebo-controlled trial with 485 obese individuals, sibutramine reduced mean weight by 4.8 kg in the 10-mg/day group and 6.1 kg in the 15-mg/day group (17).

In a 16-week double-blind trial, the addition of sibutramine, with behavioral weight counseling, to an ongoing antipsychotic regimen was evaluated in stable overweight or obese outpatients with schizophrenia (21). Twenty-one

---

**TABLE 1. Baseline Characteristics of Patients With Schizophrenia in a Placebo-Controlled Trial of Sibutramine for Olanzapine-Associated Weight Gain**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Placebo (N=18)</th>
<th>Sibutramine (N=19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>11 (61)</td>
<td>12 (63)</td>
</tr>
<tr>
<td>Female</td>
<td>7 (39)</td>
<td>7 (37)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>17 (94)</td>
<td>17 (89)</td>
</tr>
<tr>
<td>Married</td>
<td>1 (6)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>5 (28)</td>
<td>5 (26)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>12 (67)</td>
<td>13 (68)</td>
</tr>
<tr>
<td>East/South East Asian</td>
<td>0 (0)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (6)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizoaffective disorder</td>
<td>5 (28)</td>
<td>5 (26)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>13 (72)</td>
<td>14 (74)</td>
</tr>
</tbody>
</table>

| Age (years)                    | 40.7 (9.9)    | 43.2 (10.6)        |
| Educational level (years)      | 12.6 (1.7)    | 12.8 (2.1)         |
| Weight (lb)                    | 240.3 (40.1)  | 226.5 (34.6)       |
| Body mass index (kg/m²)        | 38.4 (9.1)    | 34.3 (5.8)         |
| Estimated weight gain from olanzapine (lb) | 54.2 (47.6) | 47.0 (29.9) |
| Lowest weight maintained for at least 1 year since age 21 (lb) | 158.3 (30.2) | 157.8 (31.6) |
| Blood pressure (mm Hg)         |               |                    |
| Systolic                       | 128.9 (10.8)  | 127.9 (15.4)       |
| Diastolic                      | 87.4 (6.8)    | 87.3 (13.1)        |
patients were assigned to either sibutramine (up to 15 mg/day) or placebo in a 2:1 ratio. Nineteen completed at least 4 weeks of double-blind treatment (14 taking sibutramine, five taking placebo), and 11 completed the full 16 weeks. There were no significant differences between groups on mean loss of weight (sibutramine, –8.2 lb; placebo, –9.2 lb) or body mass index (sibutramine, –1.32 kg/m²; placebo, –1.46 kg/m²) in a last-observation-carried-forward analysis. The small number of subjects and high dropout rate may have limited the opportunity to detect differences between groups.

While generally well tolerated, sibutramine can affect heart rate (average increase, 3–4 beats/minute) and blood pressure (average increase, 2 mm Hg) (17). The adverse effects of sibutramine in a phase I trial, in which participants received up to 30 mg/day, included insomnia, rapid heart rate, and diastolic hypertension. For participants receiving up to 20 mg/day, nausea, insomnia, dry mouth, rhinitis, and constipation were noted (18, 19).

We conducted a 12-week double-blind, placebo-controlled trial, adding sibutramine or placebo to the medication regimens of schizophrenia subjects with olanzapine-associated weight gain. We chose olanzapine-treated subjects for this study because weight gain is often the most problematic side effect of this otherwise well-tolerated and effective agent.

**Method**

**Subjects**

This 12-week double-blind, placebo-controlled, randomized trial was conducted in the adult outpatient clinic of an urban mental health center. The study was approved by the institutional review board of the Massachusetts Department of Mental Health. Subjects were assessed by the investigators for capacity to provide informed consent. All participants received a full explanation of the nature of the study and were required to take a quiz to ensure
Am J Psychiatry 162:5, May 2005
HENDERSON, COPELAND, DALEY, ET AL.

for Patients Receiving Placebo (N=18) or Sibutramine (N=9) for Olanzapine-Associated Weight Gain

<table>
<thead>
<tr>
<th>Week 4</th>
<th>Week 8</th>
<th>Week 12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Placebo</td>
<td>Sibutramine</td>
</tr>
<tr>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>238.1</td>
<td>39.7</td>
<td>213.9</td>
</tr>
<tr>
<td>36.8</td>
<td>6.4</td>
<td>31.8</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>115.6</td>
<td>12.3</td>
<td>104.9</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>123.1</td>
<td>13.8</td>
<td>119.1</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>1.068</td>
<td>0.090</td>
<td>1.072</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>123.9</td>
<td>16.6</td>
<td>111.5</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>41.5</td>
<td>4.7</td>
<td>39.9</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>38.0</td>
<td>3.9</td>
<td>38.1</td>
</tr>
</tbody>
</table>

Mean and SD

<table>
<thead>
<tr>
<th>Week 4</th>
<th>Week 8</th>
<th>Week 12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Placebo</td>
<td>Sibutramine</td>
</tr>
<tr>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>125.1</td>
<td>9.2</td>
<td>125.1</td>
</tr>
<tr>
<td>82.5</td>
<td>7.1</td>
<td>86.6</td>
</tr>
<tr>
<td>80.4</td>
<td>14.2</td>
<td>81.8</td>
</tr>
<tr>
<td>0.36</td>
<td>0.03</td>
<td>0.37</td>
</tr>
<tr>
<td>0.39</td>
<td>0.03</td>
<td>0.42</td>
</tr>
<tr>
<td>0.09</td>
<td>0.01</td>
<td>0.14</td>
</tr>
<tr>
<td>75.7</td>
<td>12.0</td>
<td>78.9</td>
</tr>
<tr>
<td>121.5</td>
<td>13.5</td>
<td>131.4b</td>
</tr>
<tr>
<td>84.5</td>
<td>9.8</td>
<td>90.6</td>
</tr>
<tr>
<td>81.2</td>
<td>12.0</td>
<td>84.1</td>
</tr>
<tr>
<td>0.37</td>
<td>0.03</td>
<td>0.37</td>
</tr>
<tr>
<td>0.42</td>
<td>0.03</td>
<td>0.41</td>
</tr>
<tr>
<td>0.15</td>
<td>0.23</td>
<td>0.09</td>
</tr>
<tr>
<td>79.4</td>
<td>17.2</td>
<td>79.2</td>
</tr>
</tbody>
</table>

Subjects were excluded from the study if they were unable to provide informed consent, had current substance abuse or significant medical illness (including hepatic or renal disease), had untreated hypertension, had a history of intolerance of sibutramine, were pregnant or breast-feeding, were receiving treatment with agents that induce weight loss, had a history of glaucoma, had heart disease or an abnormal electrocardiogram, or were being treated with antimigraine agents containing serotonin agonists. The patient’s clinic chart was reviewed, the preolanzapine weight was recorded, and the patient’s clinician was consulted to validate the diagnostic assessment. The subjects’ prestudy medication regimen was maintained throughout the 12-week trial. After completing baseline assessments, the subjects were randomly assigned by a research pharmacist to sibutramine and placebo.

Procedure

Each subject was given a 1-week supply of 5-mg sibutramine capsules or identical-appearing placebo capsules and instructed to take two capsules daily. Research psychiatrists (D.C.H., D.C.G., P.M.L.) were allowed to decrease the dose of the study medication to one capsule a day, as indicated for intolerable side effects. After 4 weeks, the dose was increased to three capsules, as tolerated.

Olanzapine-treated patients with a history of weight gain associated with the drug and with the diagnosis of schizophrenia or schizoaffective disorder were recruited for the study. Patients were included in the study if they met the following criteria: age of 18–65 years, well-established compliance with outpatient medication regimen, stable dose of olanzapine for at least 4 months, and no exposure to tricyclic, selective serotonin reuptake inhibitor (SSRI), or monoamine oxidase inhibitor antidepressants for 1 month. Additionally, the patients were required to have a body mass index of ≥30 kg/m² or a body mass index of 27 kg/m² plus another risk factor for cardiovascular disease (hypertension, lipid abnormality, diabetes mellitus).

Assessments

Anthropometric, nutrition, and activity assessments. Weight loss was the principal measure of efficacy. Each subject’s body weight and body mass index were measured at baseline, weekly for the first 8 weeks of the study, and at week 12. Waist circumference was measured at baseline, week 8, and week 12. The waist-to-hip ratio was calculated as the narrowest waist measurement relative to the widest hip circumference to provide information regarding body fat distribution (24–26). The percentage of body fat was calculated from skinfold measurements of the biceps, triceps, suprailiac region, and subscapular region (27, 28).

Overall behavior was assessed with a survey of food intake and dietary habits weekly for the first 8 weeks and at week 12. The subjects’ adherence to and success with individual dietary goals were assessed weekly for the first 8 weeks of the study in the behavioral nutrition group meetings and at termination, by using a 4-day food record. The food frequency questionnaire was administered at baseline and week 12. Energy and nutrient intake was analyzed by using an extensive nutrient database (Minnesota Nutrient Data System) (29, 30).

A quantitative activity questionnaire, the Modifiable Activity Questionnaire (31), was used to assess both leisure and occupational activity components. The questionnaire is administered by an interviewer and assesses activity for the past 12 months. The

http://ajp.psychiatryonline.org
A sequential testing procedure was used, and the difference at 12 weeks was tested first. The week 8 difference was tested only if the week 12 difference was significant, and the week 4 difference was tested only if the week 8 difference was significant.

To assess the possible correlations of baseline covariates with changes in weight and body mass index in the sibutramine-treated group, the slopes of weight and body mass index over time were calculated for each subject. Analyses of variance were used to test the associations of the slopes with categorical baseline covariates, while Pearson correlations were used to test the associations with continuous baseline covariates.

Results

Enrollment and Demographic Characteristics

Fifty subjects were screened, and 43 were enrolled (Figure 1). Six of these subjects were not assigned to treatments for the following reasons: abnormal ECG (N=1), lack of commitment to the study (N=4), and loss of contact (N=1). Thirty-seven subjects were randomly assigned to either sibutramine (N=19) or placebo (N=18). In all, six subjects (16%) did not complete the study. The proportions of patients who withdrew from the study were similar for sibutramine (16%) and placebo (17%). The reasons for dropping out of the study are summarized in Figure 1.

The demographic data are summarized in Table 1. There were no significant differences between the two groups with regard to gender, age, race or ethnicity, marital status, and diagnosis. Weight, body mass index, estimated weight gain from olanzapine, lowest weight since the age of 21, and systolic and diastolic blood pressures did not differ significantly between the sibutramine group and the placebo group at baseline.
butramine group exhibited significantly greater weight loss at week 12 (t = 3.22, df = 25, p = 0.004) and the increase in waist measurement (t = 3.22, df = 25, p = 0.004) and the increase in waist:hip ratio (t = 2.93, df = 25, p = 0.007) were significantly greater at week 12 in the sibutramine group than in the placebo group. Changes in measurements of abdominal circumference, neck circumference, and percentage body fat did not differ between groups. Changes in body mass index was significantly greater for the sibutramine group than for the placebo group at week 8 (t = 2.47, df = 51, p = 0.02) and week 12 (t = 2.42, df = 13, p = 0.03). Although not statistically significant, both groups demonstrated reductions in total calories. The sibutramine group showed a significant decrease in total grams of polyunsaturated fatty acids at week 12 (t = 2.31, df = 13, p = 0.04) and in the percentage of calories from polyunsaturated fatty acids at week 12 (t = 2.42, df = 13, p = 0.03).

Throughout the study, there were few clinically significant changes in hematological and biochemical laboratory parameters (Table 3). However, the mean decrease in hemoglobin A1C was significantly greater in the sibutramine group at week 12 (t = 2.93, df = 20, p = 0.008), with the baseline level controlled, suggesting an improvement in glucose metabolism. There were no differences between the groups at week 12 in measurements of random glucose, uric acid, cortisol, liver functioning, blood urea nitrogen, creatinine, thyroid functioning, or lipids. Additionally, olanzapine blood levels did not change significantly in either group during the course of the study.

### Anthropometric and Laboratory Assessments

Table 2 shows the change in weight for both groups. No significant difference in weight loss was found between the two groups at weeks 4 and 8. However, at week 12 the sibutramine group exhibited significantly greater weight loss than the placebo group (t = 2.73, df = 52, p = 0.009) (Figure 2). Reduction of body mass index was significantly greater for the sibutramine group than for the placebo group at week 8 (t = 2.47, df = 51, p = 0.02) and week 12 (t = 3.88, df = 51, p = 0.0003) (Figure 2). In addition, the decrease in waist measurement (t = 3.22, df = 25, p = 0.004) and the increase in waist:hip ratio (t = 2.93, df = 25, p = 0.007) were significantly greater at week 12 in the sibutramine group than in the placebo group. Changes in measurements of abdominal circumference, neck circumference, and percentage body fat did not differ between groups. Changes in body mass index and weight did not correlate with age, dose of olanzapine, historic lowest weight, baseline weight, weight gain with olanzapine, age at onset of schizophrenia, or educational level.

Throughout the study, there were few clinically significant changes in hematological and biochemical laboratory parameters (Table 3). However, the mean decrease in hemoglobin A1C was significantly greater in the sibutramine group at week 12 (t = 2.93, df = 20, p = 0.008), with the baseline level controlled, suggesting an improvement in glucose metabolism. There were no differences between the groups at week 12 in measurements of random glucose, uric acid, cortisol, liver functioning, blood urea nitrogen, creatinine, thyroid functioning, or lipids. Additionally, olanzapine blood levels did not change significantly in either group during the course of the study.

### Clinical Psychiatric Ratings

Overall, no significant differences between groups were found on the Positive and Negative Syndrome Scale total score and subscale scores, except at week 12, when the placebo group had a lower score than the sibutramine group on the negative symptoms subscale (t = 2.04, df = 43, p = 0.05). However, the difference in negative symptoms was not evident on the SANS total score. There were no differences between the groups with respect to the AIMS, the GAS, and the Hillside Akathisia Scale (Table 4).

### Dietary Intake

The food frequency questionnaire and the 4-day food record revealed few differences between the two groups in dietary intake from baseline to week 12 (Table 5). Although not statistically significant, both groups demonstrated reductions in total calories. The sibutramine group showed a significant decrease in total grams of polyunsaturated fatty acids at week 12 (t = 2.31, df = 13, p = 0.04) and in the percentage of calories from polyunsaturated fatty acids at week 12 (t = 2.42, df = 13, p = 0.03).

Although the difference was not statistically significant, nighttime snacking decreased in the sibutramine group, and both groups reported reductions in “overeating” that persisted throughout the study. Additionally, both groups reported reductions in eating “empty calorie” foods (cookies, chips, ice cream, french fries, cake, and pie), total calories, calories from fat, calories from sweets, fat servings, and intake (grams) of carbohydrates, total fat, and saturated fat. The percentage of calories from protein at week 12, with the baseline level controlled, was significantly increased in the placebo group (t = 2.29, df = 21, p = 0.03). Although not statistically significant, the sibutramine group also reported a reduction in meat servings. Vegetable servings increased significantly, however, when the analysis controlled for baseline, in the placebo group (t = 2.13, df = 21, p = 0.05).

---

**TABLE 4. Clinical Ratings of Patients Receiving Placebo (N=18) or Sibutramine (N=19) for Olanzapine-Associated Weight Gain**

<table>
<thead>
<tr>
<th>Rating</th>
<th>Placebo (Mean SD)</th>
<th>Sibutramine (Mean SD)</th>
<th>Placebo (Mean SD)</th>
<th>Sibutramine (Mean SD)</th>
<th>Placebo (Mean SD)</th>
<th>Sibutramine (Mean SD)</th>
<th>Placebo (Mean SD)</th>
<th>Sibutramine (Mean SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal Involuntary Movement Scale (37)</td>
<td>1.9 2.2 1.8 1.8</td>
<td>1.3 1.7 1.9 1.6</td>
<td>1.2 1.1 2.2 2.2</td>
<td>0.9 1.2 2.2 1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Assessment Scale (35) score</td>
<td>57.9 8.2 59.7 7.0</td>
<td>60.4 10.3 59.6 9.2</td>
<td>61.8 7.8 61.1 7.4</td>
<td>61.5 7.5 60.0 7.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hillside Akathisia Scale (36) score</td>
<td>1.8 3.3 0.2 1.0</td>
<td>0.4 0.9 0.3 1.1</td>
<td>0.7 1.7 1.1 2.1</td>
<td>0.8 1.8 1.1 2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive and Negative Syndrome Scale (32, 33) scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>57.4 13.6 54.9 7.1</td>
<td>53.7 10.5 54.9 7.1</td>
<td>56.8 12.4 53.2 7.3</td>
<td>51.9 11.2 53.0 4.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative symptoms</td>
<td>15.9 5.4 15.4 4.0</td>
<td>16.7 6.9 14.6 5.2</td>
<td>15.5 5.9 15.6 4.7</td>
<td>13.4 4.5 16.0 3.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive symptoms</td>
<td>12.9 4.8 13.8 4.2</td>
<td>11.6 3.8 13.5 4.0</td>
<td>13.2 4.6 12.8 3.3</td>
<td>12.2 5.6 12.3 2.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive-negative</td>
<td>28.8 8.0 28.8 4.2</td>
<td>28.4 7.2 28.1 6.3</td>
<td>28.8 8.1 28.4 5.5</td>
<td>25.6 7.6 28.3 3.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>10.1 3.0 9.7 3.8</td>
<td>7.0 4.5 8.1 4.9</td>
<td>7.1 4.7 6.7 4.7</td>
<td>7.3 5.1 6.3 4.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale for the Assessment of Negative Symptoms (34) score</td>
<td>34.3 16.7 29.1 10.3</td>
<td>35.9 15.7 32.0 15.0</td>
<td>33.1 10.4 28.8 16.8</td>
<td>31.9 11.8 34.4 9.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Numbers of subjects vary because of different patterns of missing data.
b Significant difference between groups in change from baseline, controlled for baseline levels (p < 0.05, two-tailed t test).
## Adverse Events

Diastolic blood pressure, heart rate, and most ECG variables did not differ significantly between groups throughout the study. Systolic blood pressure increased significantly in the sibutramine group at week 12 ($t=2.23$, df=51, $p=0.04$). Dry mouth, constipation, blurred vision, excessive thirst (anticholinergic side effects), interrupted sleep, and shortened sleep increased in more than 5% of the subjects in the sibutramine group, and increases in these conditions occurred in at least twice as many patients in the sibutramine group as in the placebo group. On the other hand, dizziness increased to a greater degree in the placebo group (Table 6). No subjects were withdrawn from the study because of adverse events.

## Three-Month Follow-Up

Ten placebo subjects and 12 sibutramine subjects returned 3 months after the end of the study for a follow-up assessment. While body mass index, weight, waist circumference, and waist-hip ratio had decreased significantly at week 12, the change from baseline of the sibutramine group was not statistically different from that of the placebo group for waist circumference (sibutramine: mean=−5.3 cm, SD=8.0; placebo: mean=−2.3 cm, SD=16.7), waist-hip ratio (sibutramine: mean=−0.03, SD=0.05; placebo: mean=−0.1, SD=0.2), weight (sibutramine: mean=−5.7 lb, SD=14.9; placebo: mean=−0.3 lb, SD=15.8), or body mass index (sibutramine: mean=−0.2 kg/m², SD=2.9; placebo: mean=−0.04 kg/m², SD=2.5) at the time of the 3-month follow-up.

### TABLE 5. Dietary Assessments of Patients Receiving Placebo or Sibutramine for Olanzapine-Associated Weight Gain

<table>
<thead>
<tr>
<th>Assessment of Daily Diet</th>
<th>Placebo</th>
<th>Sibutramine</th>
<th>Placebo</th>
<th>Sibutramine</th>
<th>Placebo</th>
<th>Sibutramine</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total calories (kcal)</td>
<td>1912.9</td>
<td>869.7</td>
<td>1784.7</td>
<td>743.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of calories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>47.5</td>
<td>8.1</td>
<td>47.1</td>
<td>10.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>15.5</td>
<td>2.7</td>
<td>14.4</td>
<td>3.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat</td>
<td>38.1</td>
<td>7.8</td>
<td>39.6</td>
<td>8.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweets</td>
<td>18.4</td>
<td>14.1</td>
<td>22.7</td>
<td>12.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total calories (kcal)</td>
<td>1575.8</td>
<td>393.3</td>
<td>1355.8</td>
<td>798.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of calories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>180.7</td>
<td>49.5</td>
<td>162.5</td>
<td>39.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>68.9</td>
<td>18.6</td>
<td>52.0</td>
<td>13.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fat</td>
<td>65.2</td>
<td>25.2</td>
<td>55.2</td>
<td>18.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fat</td>
<td>24.2</td>
<td>9.9</td>
<td>19.4</td>
<td>5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Servings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fruit</td>
<td>1.0</td>
<td>0.9</td>
<td>1.0</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vegetables</td>
<td>0.8</td>
<td>0.5</td>
<td>0.8</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grains</td>
<td>5.1</td>
<td>4.1</td>
<td>5.2</td>
<td>4.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dairy products</td>
<td>2.1</td>
<td>1.3</td>
<td>2.2</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meat</td>
<td>1.8</td>
<td>1.2</td>
<td>2.0</td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat</td>
<td>4.0</td>
<td>3.0</td>
<td>4.1</td>
<td>2.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 12</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total calories (kcal)</td>
<td>1578.6</td>
<td>471.9</td>
<td>1344.3</td>
<td>930.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of calories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>168.7</td>
<td>28.6</td>
<td>166.2</td>
<td>115.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>43.8</td>
<td>5.8</td>
<td>48.9</td>
<td>3.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fat</td>
<td>20.3</td>
<td>1.5</td>
<td>15.7</td>
<td>3.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fatty acids</td>
<td>2.0</td>
<td>2.8</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intake</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrates (g)</td>
<td>167.8</td>
<td>124.6</td>
<td>185.1</td>
<td>100.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein (g)</td>
<td>36.6</td>
<td>29.7</td>
<td>47.4</td>
<td>24.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fat (g)</td>
<td>35.7</td>
<td>4.8</td>
<td>27.6</td>
<td>8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fatty acids (g)</td>
<td>2.0</td>
<td>2.8</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fiber (g)</td>
<td>15.1</td>
<td>5.8</td>
<td>9.4</td>
<td>6.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium (mg)</td>
<td>3033.8</td>
<td>820.2</td>
<td>2278.5</td>
<td>1252.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Snacking (nights/week)</td>
<td>2.3</td>
<td>0.8</td>
<td>3.3</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a N=17 for placebo and for sibutramine.
b Significant difference between groups in change from baseline, controlled for baseline levels ($p<0.05$, two-tailed t test).
c N=16 for placebo and N=14 for sibutramine.
dex, waist circumference, and the waist-hip ratio; these changes represent a substantial reduction in cardiovascular risk factors. There was a significant improvement in hemoglobin A1C in the sibutramine group, suggesting, as expected, that weight loss in this group had a positive impact on glucose metabolism. Weight loss did not result in improvements on any lipid measurement, suggesting that other mechanisms are responsible for olanzapine-associated hyperlipidemia.

The reductions in weight, body mass index, waist circumference, and waist-hip ratio were not sustained at 3 months following the study completion. The subjects who received sibutramine gained weight and returned to near their baseline weight. This suggests that, as expected, the benefits of sibutramine do not continue after discontinuation of the drug.

In reversing olanzapine-associated weight gain, sibutramine may provide a clue as to the mechanism of olanzapine’s metabolic effects. Sibutramine acts as an agonist at the 5-HT2C receptor, and a polymorphism of the promoter region of the serotonin 5-HT2C receptor gene has been associated with antipsychotic-induced weight gain (42). Among 32 Chinese patients with first-episode schizophrenia, the 10 who had the −759C/T polymorphism showed significantly less weight gain with clozapine than those without this allele. Olanzapine acts as an antagonist at the 5-HT2C receptor. It is possible that sibutramine’s competitive agonist effect at the 5-HT2C receptor decreases an olanzapine-induced increase in appetite (43).

Following a physical examination, baseline safety assessments, and an ECG, sibutramine appeared to be safe and effective at a dose of 10 or 15 mg/day in overweight and obese schizophrenia subjects. Sibutramine was well tolerated, and no serious adverse events occurred. There were no changes in the results of safety assessments, including renal, hepatic, and thyroid tests. Olanzapine blood levels were not altered. While weight loss is usually associated with a decrease in blood pressure, we observed a mild increase in systolic blood pressure. Typical sibutramine-related side effects, including dry mouth and constipation, were observed. However, the dropout rates in the two groups were similar.

Sibutramine did not worsen psychotic symptoms. Whereas the negative symptoms subscore on the Positive and Negative Syndrome Scale decreased significantly more in the placebo group than in the sibutramine group, changes in negative symptoms were not evident on the SANS for either group.

**Limitations**

The results of this study may not be generalizable to other populations, particularly in the absence of a nutritional counseling program. Sibutramine was not administered to patients taking other serotonergic agents, such as SSRIs and migraine medications. Sibutramine should be used with caution in patients treated with SSRIs, as the risk of serotonin syndrome may be increased. It is also not clear if sibutramine would be an effective weight loss agent in patients treated with other medications that cause significant weight gain, such as clozapine. Additionally, this clinical trial was short (12 weeks), and it is not clear whether the observed weight reductions can be maintained with long-term use of sibutramine.

**Conclusions**

Sibutramine treatment improved several health status markers that are predictive of cardiovascular disease. This degree of weight loss in obese schizophrenia patients could, if sustained, substantially reduce morbidity and mortality. Sibutramine appeared to be a safe and effective weight loss agent in this study group. Clinicians should consult with primary care clinicians to determine the safety and necessity of sibutramine for individual patients. Additional studies are necessary to establish the long-term weight loss effectiveness and safety in schizophrenia patients whose weight gain from antipsychotic medications jeopardizes their cardiovascular health.

**TABLE 6. Patients Who Reported Adverse Events in a 12-Week Placebo-Controlled Trial of Sibutramine for Olanzapine-Associated Weight Gain**

<table>
<thead>
<tr>
<th>Adverse Event</th>
<th>Placebo (N=17)</th>
<th>%</th>
<th>Sibutramine (N=19)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid heart rate</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Headaches</td>
<td>5</td>
<td>29</td>
<td>7</td>
<td>37</td>
</tr>
<tr>
<td>Difficulty falling asleep</td>
<td>5</td>
<td>29</td>
<td>7</td>
<td>37</td>
</tr>
<tr>
<td>Interrupted sleepa</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>Shortened sleepa</td>
<td>2</td>
<td>12</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>Early waking</td>
<td>3</td>
<td>18</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Decreased appetite</td>
<td>11</td>
<td>65</td>
<td>12</td>
<td>63</td>
</tr>
<tr>
<td>Excessive appetite</td>
<td>6</td>
<td>35</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Excessive thirsta</td>
<td>2</td>
<td>12</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>Dry moutha</td>
<td>2</td>
<td>12</td>
<td>8</td>
<td>42</td>
</tr>
<tr>
<td>Nausea</td>
<td>6</td>
<td>35</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>Constipationa</td>
<td>2</td>
<td>12</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>Rhinitisa</td>
<td>4</td>
<td>24</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>Dizziness</td>
<td>3</td>
<td>18</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Blurred visiona</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>16</td>
</tr>
</tbody>
</table>

*Occurred in more than 5% of the patients taking sibutramine and was at least twice as common in the placebo group.*

---

Received Feb. 8, 2004; revision received April 5, 2004; accepted May 14, 2004. From the Schizophrenia Program, Weight Center, Endocrine Unit, and Biostatistics Center, Massachusetts General Hospital, Boston; and the Departments of Psychiatry and Medicine, Harvard Medical School, Boston. Address correspondence and reprint requests to Dr. Henderson, Freedom Trail Clinic, 25 Staniford St., Boston, MA 02114; dhenderson@partners.org (e-mail).

Supported by a Young Investigator Award to Dr. Henderson from the National Alliance for Research on Schizophrenia and Depression and by an investigator-initiated independent research grant from Eli Lilly and Co. Drug and placebo were provided by Knoll Pharmaceuticals (Abbott Laboratories).
References

1. Davis JM, Chen N, Glick ID: A meta-analysis of the efficacy of second-generation antipsychotics. Arch Gen Psychiatry 2003; 60:553–564
34. Andreasen NC: Scale for the Assessment of Negative Symptoms (SANS). Iowa City, University of Iowa, 1983
Does Memory of a Traumatic Event Increase the Risk for Posttraumatic Stress Disorder in Patients With Traumatic Brain Injury? A Prospective Study

Sharon Gil, Ph.D.
Yael Caspi, Ph.D.
Irit Zilberman Ben-Ari, Ph.D.
Danny Koren, Ph.D.
Ehud Klein, M.D.

Objective: The present study examined prospectively the relationship between memory of the traumatic event and subsequent development of posttraumatic stress disorder (PTSD). More specifically, the aims of this study were to 1) investigate the possibility that lack of memory of the traumatic event might be a protective factor; 2) assess whether memory of the traumatic event equally affects the three symptom clusters of PTSD: reexperiencing, avoidance, and hyperarousal; and 3) explore the predictive value of memory of the traumatic event for the development of subsequent PTSD in the immediate aftermath of the event.

Method: One hundred twenty subjects with mild traumatic brain injury who were hospitalized for observation were assessed immediately after the trauma and followed up 1 week, 3 months, and 6 months later. All participants underwent psychiatric evaluation and self-assessment of their memory of the traumatic event.

Results: Overall, 17 (14%) of the participants met full criteria for PTSD at 6 months. Subjects with memory of the traumatic event were significantly more likely to develop PTSD than those without memory of the traumatic event; the difference between the groups resulted primarily from the reexperiencing cluster. Logistic regression analysis revealed that memory of the traumatic event within the first 24 hours is a strong predictor of PTSD 6 months after the event.

Conclusions: Our study indicated that memory of a traumatic event is a strong predictor and a potential risk factor for subsequent development of PTSD. Future studies are needed to show whether these findings can be generalized to other traumatic conditions.

The intricate system of memory is commonly thought of as composed of two primary pathways. The first is referred to as explicit, or declarative, memory. This relates to conscious awareness of facts and requires focal attention for processing; it is probably mediated by the medial temporal lobe system that includes the hippocampal formation and related structures that enable verbal representation (1–3). The second pathway, referred to as implicit, or nondeclarative, memory, relates to memories acquired during skill learning, habit formation, and simple, classic conditioning. It also refers to other knowledge expressed through performance rather than recollection. These memories are believed to be less accessible to consciousness (4).

Traumatic memories share both explicit and implicit features and are processed differently than ordinary memories (5). This results in failure to organize the traumatic event into a coherent verbally represented narrative (5, 6). The abnormal nature of the traumatic memories is considered to be a central feature of posttraumatic stress disorder (PTSD). This is manifested by hypermnestic symptoms, such as reexperiencing, intrusive thoughts, nightmares, and flashbacks on one hand and, at the same time, impaired memory of the traumatic event in the form of amnesia and delayed recall, which is also a well-known phenomenon in traumatized individuals (7). This raises the question as to the extent to which each constitutes a risk factor for PTSD.

In recent years, researchers have focused on traumatic memory and the mechanisms by which it operates to better understand the risk factors associated with it in the development of PTSD. However, empirically based studies are limited by practical and ethical constraints (8) and thus, naturalistic conditions, in which traumatic memories are compromised, may serve as models to study this question. Traumatic brain injury, which is commonly associated with impaired memory, has been used as a naturally occurring model for the investigation of the relationship between memory and PTSD (9–12).

Some of the studies that focused on traumatic brain injury have provided evidence that traumatic events involving traumatic brain injury are associated with reduced prevalence of PTSD, consistent with the view that amnesia of the traumatic event may play a protective role in this regard (13, 14). Other studies, however, have shown that PTSD is fairly prevalent among patients with traumatic brain injury but is less prevalent among those with traumatic brain injury and amnesia of the traumatic event (15). This raises the question as to whether lack of memory of the traumatic event serves as a protective factor in the development of PTSD.

The present study examined prospectively the relationship between memory of the traumatic event and subsequent development of PTSD, with the following specific aims: 1) to investigate the possibility that lack of memory of the traumatic event might be a protective factor; 2) to assess whether memory of the traumatic event equally affects the three symptom clusters of PTSD: reexperiencing, avoidance, and hyperarousal; and 3) to explore the predictive value of memory of the traumatic event for the development of subsequent PTSD in the immediate aftermath of the event.

Method: A prospective study was conducted with 120 subjects with mild traumatic brain injury who were hospitalized for observation immediately after the trauma and followed up 1 week, 3 months, and 6 months later. All participants underwent psychiatric evaluation and self-assessment of their memory of the traumatic event.

Results: Overall, 17 (14%) of the participants met full criteria for PTSD at 6 months. Subjects with memory of the traumatic event were significantly more likely to develop PTSD than those without memory of the traumatic event; the difference between the groups resulted primarily from the reexperiencing cluster. Logistic regression analysis revealed that memory of the traumatic event within the first 24 hours is a strong predictor of PTSD 6 months after the event.

Conclusions: Our study indicated that memory of a traumatic event is a strong predictor and a potential risk factor for subsequent development of PTSD. Future studies are needed to show whether these findings can be generalized to other traumatic conditions.
brain injury, supporting the view that the traumatic brain injury and PTSD are not mutually exclusive (15, 16).

A significant limitation of these studies, however, is the fact that they did not directly evaluate or control for actual memory of the traumatic event. That is, the degree to which victims of traumatic brain injury, in fact, remember the traumatic event was not assessed. Although it is reasonable to assume that traumatic brain injury impairs memory of the traumatic event (9), there exists a significant variability within patients with traumatic brain injury with regard to the amount and quality of their memory of the traumatic event. It is quite possible that this variability provides an explanation for these conflicting results.

The main purpose of the present study was to overcome this limitation by direct assessment of the relationship between explicit memory of the traumatic event and subsequent development of PTSD in participants who had experienced a traumatic event associated with traumatic brain injury. More specifically, the goals of the present study were to further investigate the assumption that lack of memory of the traumatic event serves as a protective factor against the development of PTSD and to assess whether lack of memory of the traumatic event equally affects all three PTSD symptom clusters, namely, reexperiencing, avoidance, and hyperarousal. Finally, an attempt was made to assess the relative contribution of memory of the traumatic event to the development of PTSD in the context of other risk factors.

Method

Participants

The study population included 120 patients with traumatic brain injury who were recruited from two surgical wards at Rambam Medical Center to which they were admitted for medical care after traumatic brain injury. The participants had to be between the ages of 18 and 50 and be fluent in Hebrew. Excluded were those actively receiving psychiatric care at the time of injury or who had a prior history of head trauma, cognitive deficit, substance abuse, or a major untreated medical condition. After receiving a detailed description of the study, each subject provided written informed consent for participation in the study, which was approved by the institutional review board at Rambam Medical Center.

Of 198 eligible subjects, 44 (23%) refused to participate and 34 (17%) dropped out during follow-up. Thus, the final cohort consisted of 120 subjects who completed the entire follow-up. These were relatively young subjects (mean age=31.4 years, SD=2.7), with an average education of high school (mean=12.6 years, SD=2.5). The participants were predominantly men (58%), married (50%) or single (50%), Israeli-born (68%), without a prior history of physical (80%) or psychiatric (59%) disorders, who were injured primarily in traffic accidents (90%), with a mean injury severity score of 5.8 (SD=3.6), indicating mild physical injury. None of the participants was unconscious at the time of the admission, and their scores on the Glasgow Coma Scale (17) were in the upper range (score=13–15), which excludes loss of consciousness.

Procedure and Measures

The initial evaluation took place within the first 24 hours after the injury, during hospitalization. The participants were then invited for further evaluation after 7 to 10 days, 4 weeks, and 6 months from the traumatic event.

At the first interview, the participants were asked to provide information regarding their personal background and self-assessment of their memory of trauma. In addition, the severity of physical injury was assessed with the injury severity score, which is the sum of the Abbreviated Injury Scale (10), rated by a trained physician. Acute dissociative symptoms were measured with the Peritraumatic Dissociation Questionnaire (11).

Memory of the traumatic event was evaluated with a nine-item self-report questionnaire, which was specifically developed for the present study since no such tool was available during the time of the study. The questionnaire assesses the participants’ memory of the traumatic event regarding the following aspects of the trauma: 1) what was the event; 2) where did the event take place; 3) who (other than you) was involved in the event; 4) when did the event occur; 5) sights from the event; 6) sounds from the event; 7) odors from the event; 8) things you said during or after the event; and 9) things other people said during or after the event. Each of the nine items is rated on a 4-point Likert scale, ranging from 1 (no memory) to 4 (good memory), and a total score is derived by calculating a mean score for the memory of the nine traumatic event questions. The questionnaire, which was administered at all four time points, was found to be reliable, with a Cronbach’s alpha of 0.91. In follow-up assessments, the respondents were instructed to complete the Memory of Traumatic Event Questionnaire with an attempt to disregard their previous responses.

PTSD symptoms were assessed with two instruments: the Clinician-Administered PTSD Scale (12) and the Posttraumatic Stress Scale (18). The Clinician-Administered PTSD Scale is a 17-item questionnaire administered by a trained clinician to assess levels of posttraumatic symptoms during the previous 2 weeks. The severity of each item is rated on a 5-point Likert scale ranging from 0 (not at all) to 4 (very much). The total severity score, ranging from 0 to 68, is calculated as the mean of participants’ ratings on the 17 items.

The Posttraumatic Stress Scale is a 17-item self-report questionnaire assessing levels of posttraumatic symptoms during the previous 2 weeks. The severity of each item is rated on a 4-point Likert scale ranging from 0 (not at all) to 3 (very much). The total severity score, ranging from 0 to 51, is calculated as the mean of participants’ ratings on the 17 items.

In addition, three subscales were calculated corresponding to the definition of the diagnosis of PTSD: reexperiencing (items 1–5), avoidance (items 6–12), hyperarousal (items 13–17) (on both the Clinician-Administered PTSD Scale and the Posttraumatic Stress Scale). Pearson’s correlation analysis showed that the total score of each scale is highly correlated with its subscales (Clinician-Administered PTSD Scale: reexperiencing: r=0.89, avoidance: r=0.81, hyperarousal: r=0.88; Posttraumatic Stress Scale—reexperiencing: r=0.88, avoidance: r=0.82, hyperarousal: r=0.80).

Depression and anxiety symptoms were assessed with the Beck Depression Inventory (19) and the Beck Anxiety Inventory (20), respectively. All ratings were carried out at 7 to 10 days, 4 weeks, and 6 months.

In order to determine lifetime and current diagnoses of any DSM-IV axis I disorder (major psychiatric disorder), the Structured Clinical Interview for DSM-IV, Nonpatient Edition (21), was administered at 1 week and 6 months, respectively. The interview was carried out by an experienced and specially trained clinical social worker.

To detect general cognitive deficits that might have resulted from traumatic brain injury but were unrelated specifically to the memory of the traumatic event, the Automated Neuropsychological Assessment Metrics (22) was administered at the end of the study. The Automated Neuropsychological Assessment Metrics includes a battery of standardized tests designed for clinical use,
TABLE 1. Demographic Characteristics of Subjects With and Without Posttraumatic Stress Disorder (PTSD) and With and Without Memory of a Traumatic Event

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Subjects With PTSD (N=17)</th>
<th>Subjects Without PTSD (N=103)</th>
<th>Analysis</th>
<th>Subjects With Memory of a Traumatic Event (N=55)</th>
<th>Subjects Without Memory of a Traumatic Event (N=65)</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>t (df=118)</td>
<td>Mean</td>
</tr>
<tr>
<td>Age (years)</td>
<td>33.8</td>
<td>3.1</td>
<td>31.0</td>
<td>2.3</td>
<td>2.9*</td>
<td>32.7</td>
</tr>
<tr>
<td>Education (years)</td>
<td>12.3</td>
<td>2.4</td>
<td>12.5</td>
<td>2.7</td>
<td>n.s.</td>
<td>12.1</td>
</tr>
<tr>
<td>Injury severity score</td>
<td>6.0</td>
<td>3.9</td>
<td>5.8</td>
<td>3.3</td>
<td>n.s.</td>
<td>5.9</td>
</tr>
<tr>
<td>Gender</td>
<td>n.s.</td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
</tr>
<tr>
<td>Male</td>
<td>11</td>
<td>64</td>
<td>59</td>
<td>53</td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
<td>36</td>
<td>44</td>
<td>43</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Marital status</td>
<td>n.s.</td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
</tr>
<tr>
<td>Married</td>
<td>9</td>
<td>52</td>
<td>51</td>
<td>49</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Single</td>
<td>8</td>
<td>48</td>
<td>52</td>
<td>51</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Country of origin</td>
<td>n.s.</td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
</tr>
<tr>
<td>Israel</td>
<td>11</td>
<td>64</td>
<td>71</td>
<td>68</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>Not Israel</td>
<td>6</td>
<td>36</td>
<td>32</td>
<td>32</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>History of physical injury</td>
<td>n.s.</td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
</tr>
<tr>
<td>Yes</td>
<td>4</td>
<td>23</td>
<td>20</td>
<td>20</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>No</td>
<td>13</td>
<td>77</td>
<td>83</td>
<td>80</td>
<td>23.1**</td>
<td>42</td>
</tr>
<tr>
<td>History of psychiatric disorder</td>
<td>24</td>
<td>44</td>
<td>25</td>
<td>38</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>70</td>
<td>37</td>
<td>35</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>30</td>
<td>66</td>
<td>65</td>
<td></td>
<td>31</td>
</tr>
</tbody>
</table>

*p<0.05.  **p<0.001.

from which a representative subset of the following six tests was chosen for the present study: 1) the Stanford Sleepiness Scale, 2) the Simple Reaction Time, 3) the Memory Search, 4) the Code Substitution, 5) the Matching to Sample, and 6) the Running Memory Continuous Performance Test.

Data Analysis

First, the distribution of PTSD and the memory of the traumatic event were each calculated separately at the different time points. The relationship between PTSD and the memory of the traumatic event was then assessed on both categorical and ordinal levels by using chi-square analysis (with Yates’s correction) and multivariate repeated-measures analysis of variance (MANOVA), respectively.

In order to explore the differential effect of memory of the traumatic event on the three PTSD symptom clusters, an additional MANOVA with repeated measures was performed by using memory of the traumatic event as an independent variable and the three PTSD symptom clusters (i.e., reexperiencing, avoidance, hyperarousal) as dependent variables.

Automated Neuropsychological Assessment Metrics results were analyzed with two-tailed Student’s t tests, comparing subjects with and without PTSD and those with and without memory of the traumatic event.

Finally, a logistic regression analysis was performed, with PTSD status at 6 months as the dependent variable. Only the variables that showed a significant association with PTSD in previous analyses were used as independent variables in the regression model.

Results

Prevalence and Longitudinal Course of PTSD

Six months after the traumatic event, 14% (17 of 120) of the participants met diagnostic criteria for PTSD and 10% (12 of 120) met diagnostic criteria for psychiatric disorder other than PTSD. Table 1 presents the demographic characteristics of the PTSD and the non-PTSD groups and shows that they were similar on all parameters except age (t=2.9, df=118, p<0.05), which was found to be higher in the PTSD group. Figure 1 presents the longitudinal course of posttraumatic symptoms in the two groups, as measured by the Clinician-Administered PTSD Scale and the Posttraumatic Stress Scale. Since the Clinician-Administered PTSD Scale and the Posttraumatic Stress Scale are correlated, it would have been possible to analyze them jointly. However, because most studies employ only one of these instruments, separate analyses enable comparisons across the different study groups. In addition, results from a joint MANOVA showed similar results to those reported. Visual inspection of Figure 1 reveals that an initial, relatively small difference between the groups appears to have increased progressively over the 6-month follow-up period. In line with this impression, MANOVA with repeated measures showed a significant group-by-time interaction for the total score of the Clinician-Administered PTSD Scale (Wilks’s lambda: F=5.1, df=4, 69, p<0.001) and the Posttraumatic Stress Scale (Wilks’s lambda: F=4.7, df=4, 69, p<0.001). The main contribution to the overall interaction effect of PTSD on the Clinician-Administered PTSD Scale and on the Posttraumatic Stress Scale came from the second time interval, namely, the period from 1 week to 1 month (Clinician-Administered PTSD Scale: F(time)=5.1, df=1, 118, p<0.001, and the Posttraumatic Stress Scale: F(time)=4.1, df=1, 118, p<0.001). Duncan’s post hoc test revealed that the differences between the PTSD and the non-PTSD groups on both scales were significant at 1 month (Clinician-Administered PTSD Scale: F=3.1, df=1,
118, p<0.01, and the Posttraumatic Stress Scale: F=3.4, df=1, 118, p<0.001) but not at 1 week.

**Memory of the Traumatic Event**

Figure 2 presents the distribution of respondents on the Memory of Traumatic Event Questionnaire at the initial interview, namely, 24 hours after the injury. As shown, a bimodal distribution is evident, with most participants reporting either good memory of the traumatic event or total lack of memory of the traumatic event. Consequently, a categorical approach was taken, with participants divided into two groups with the median (2.1) of the ordinal scale used as the cutoff point. Thus, the subjects were referred to as either “having memory of the traumatic event” or “having no memory of the traumatic event.” Accordingly, 55 participants (45%) had memory of the traumatic event and 65 (55%) had no memory of the traumatic event.

As shown in Table 1, the demographic characteristics of the participants with and without memory of the traumatic event were found to be comparable on all measures except gender (χ²=16.7, df=118, p<0.001).

Intraclass correlations showed stability of memory of the traumatic event over time, with a strong correlation between the scores on the questionnaire at the different time points (F=3.95, df=1, 118, p<0.001; alpha=0.76).

**Relationship Between Memory of Traumatic Event and PTSD**

Examination of the relationship between memory of the traumatic event and PTSD at 6 months revealed that PTSD was significantly more prevalent among participants with memory of the traumatic event than among those without memory of the traumatic event (χ²=20.1, df=118, p<0.001). Among the 55 participants with memory of the traumatic event, 13 (23%) had PTSD, whereas only four (6%) of the 65 participants without memory of the traumatic event developed PTSD. Thus, the crude relative risk for PTSD among the participants with memory of the traumatic event was almost five times higher than among those without memory of the traumatic event (odds ratio=4.6, confidence interval [CI]=1.1–9.9, p<0.001).

MANOVAs with repeated measures indicated that the association between memory of the traumatic event and PTSD remained stable over 6 months, as shown by significant effects on PTSD symptoms of both memory of the traumatic event (Clinician-Administered PTSD Scale: F=4.7, df=1, 118, p<0.001, and Posttraumatic Stress Scale: F=4.9, df=1, 118, p<0.001) and time (Clinician-Administered PTSD Scale: F=4.9, df=1, 118, p<0.001, and Posttraumatic Stress Scale: F=4.6, df=1, 118, p<0.001). No group-by-time interaction was found; namely, both the Clinician-Administered PTSD Scale and Posttraumatic Stress Scale scores were significantly higher among the participants with memory of the traumatic event than among those without memory of the traumatic event at all time points.

Next we assessed the differential effect of memory of the traumatic event on each of the three PTSD clusters (reexperiencing, avoidance, and hyperarousal) using MANOVAs with repeated measures. This analysis revealed that the difference between the groups was primarily due to the reexperiencing symptom cluster for both the Clinician-Administered PTSD Scale (F≥3.4, df=1, 118, p<0.01). No significant differences were found on the avoidance and hyperarousal symptom clusters.

Analysis of the cognitive performance tests (Automated Neuropsychological Assessment Metrics) showed that no significant differences were evident between participants with and without memory of the traumatic event or be-
tween participants with and without PTSD on any of the measures.

**Logistic Regression Analysis**

In the final stage, a logistic regression analysis was performed, with PTSD status at 6 months as the dependent variable. The independent variables included in this model were only those that had shown significant associations with PTSD in previous analyses: memory of the traumatic event (within 24 hours after the event), dissociative reaction (within 24 hours after the event), acute PTSD symptoms (within 1 week after the event on both the Clinician-Administered PTSD Scale and the Posttraumatic Stress Scale), depressive symptoms (within 1 week after the event), anxiety symptoms (within 1 week after the event), age, history of psychiatric disorder, and gender.

The results indicated that accounting for all of these correlates and memory of the traumatic event was strongly associated with PTSD at 6 months. Thus, respondents with memory of the traumatic event were more than twice as likely to have PTSD compared with those without memory of the traumatic event (odds ratio=2.2, CI=1.0–10.1). In addition, acute posttraumatic symptoms (Clinician-Administered PTSD Scale: odds ratio=5.3, CI=1.1–9.3, and Posttraumatic Stress Scale: odds ratio=5.2, CI=1.0–9.4), depressive symptoms (odds ratio=5.1, CI=1.0–9.2), anxiety symptoms (odds ratio=4.9, CI=1.0–9.1) within 1 week of the traumatic event as well as a history of psychiatric disorder (odds ratio=3.7, CI=1.1–8.9) were all associated with an increased risk for PTSD at 6 months. The overall model explained 38% of the variance (Nagelkerke R²=38, p<0.001).

**Discussion**

The main goal of the present study was to assess the relationship between memory of the traumatic event and the subsequent development of PTSD in subjects with traumatic brain injury with a prospective design.

The central finding of our study is that memory of a traumatic event is positively associated with the risk for development of PTSD, while lack of memory of a traumatic event decreases the risk and might, in fact, play a protective role. Thus, along with other factors, such as history of previous trauma (23), previous psychiatric morbidity (23, 24), and physical injury (24, 25), memory of a traumatic event appears to be another risk factor for PTSD. Moreover, memory of a traumatic event assessed as early as 24 hours posttrauma appears to be a strong predictor of PTSD at 6 months.

A closer examination of the specific effect of memory of a traumatic event on the three PTSD symptom clusters reveals that it affects mainly the reexperiencing symptom cluster while the avoidance and hyperarousal clusters did not differ between those with and without memory of a traumatic event. This differential effect is not surprising given the role of memory in the reexperiencing symptom cluster. Nevertheless, this finding is different from our previous results in studies with injured survivors of motor vehicle accidents (24) and with combat survivors (25) in which bodily injury, another risk factor for PTSD, affects all three PTSD symptoms similarly. This difference might suggest that PTSD is not a homogeneous entity but rather a multidimensional and complex disorder.

Our findings further highlight the predictive value of memory of the traumatic event. Given the stability of memory of the traumatic event over the first 6 months after the trauma, memory of the traumatic event reported within 24 hours after the event appears to be a strong predictor of the presence of PTSD 6 months later. Thus, by merely asking traumatic brain injury survivors immediately after the first 24 hours after the event whether they remember its details may identify those who are at risk for the development of PTSD and thus in need of therapeutic intervention.

These findings seem to be in contrast with the theoretical assumptions underlying many of the therapeutic interventions with patients suffering from PTSD (e.g., exposure, abreaction, hypnosis) that highlight the importance of eliciting traumatic memories as part of the recovery process (26). Our findings indicate that at least for traumatic brain injury survivors without memory of the traumatic event, forgetting may be protective, in which case the process of deliberate recollection and remembering may be harmful rather than therapeutic.

It is noteworthy, however, that 6% of the participants in this study without memory of the traumatic event did meet

---

**FIGURE 2. Distribution of 120 Subjects With and Without Posttraumatic Stress Disorder on the Memory of Traumatic Event Questionnaire**

*The questionnaire was developed for the study and assesses memory of a traumatic event regarding the following aspects of the trauma: 1) what was the event; 2) where did the event take place; 3) who (other than you) was involved in the event; 4) when did the event occur; 5) sights from the event; 6) sounds from the event; 7) odors from the event; 8) things you said during or after the event; and 9) things other people said during or after the event. The percent of subjects who scored 0 (no memory) or 1–4 (good memory) on the Memory of Traumatic Event Questionnaire are shown in the figure.*

---

full diagnostic criteria for PTSD at 6 months, indicating that PTSD can develop in the absence of memory of the traumatic event. One possible mechanism by which these results could be explained is that emotionally charged traumatic memories are initially processed with brain circuits that bypass cortical structures and are mediated primarily through the amygdala, resulting in the formation of implicit (unconscious) memories (27–29). In addition, the stress-induced secretion of glucocorticosteroids, which have been shown to impair hippocampal functioning, may disrupt the formation of explicit memory (30).

The generalizability of these findings beyond traumatic brain injury to other conditions, which are associated with impaired or reduced memory of the traumatic event, is at this point speculative and needs to be addressed cautiously. However, in a study of Holocaust survivors (31), it was shown that decrease in dream recall serves as a defensive adaptive function. Taking this even further, with all due caution, one might question whether “deliberate” disruption of memory of the traumatic event might prove therapeutically beneficial. This possibility was recently addressed in a double-blind study that examined the severity of acute PTSD symptoms among 18 subjects who were given 40 mg of propranolol 6 hours after trauma in comparison with the severity of symptoms among 23 participants who received placebo (32). Results showed that participants in the experimental group tended to exhibit lower levels of PTSD symptoms 10 days after the traumatic event. If further corroborated, these findings may support the notion that not only does lack of memory of the traumatic event protect against the development of PTSD but also the pharmacologically induced disruption of the consolidation of traumatic memories can be therapeutically beneficial for trauma survivors.

Several limitations of our study deserve attention. First, participants in this study were asked only to rate the degree to which they believe they did or did not remember certain aspects of the trauma. Thus, our instrument did not actually assess memory of the traumatic event but, rather, one’s confidence in the memory for details of the trauma. This distinction is important because our findings cannot rule out the possibility that false memory of the traumatic event can also increase the risk for PTSD or, alternatively, that just lack of confidence in one’s own memory can also serve as a protective factor against PTSD. Nevertheless, the finding that appraisal of memory of the traumatic event, regardless of its objective accuracy, appears to be a strong predictor of subsequent PTSD, is potentially important, and has both theoretical and clinical implications. A Traumatic Memory Inventory (8), which was not available at the time of our study, may prove helpful in further clarifying this in future studies. Second, we asked our participants only about their memory of the traumatic event but not of other nontraumatic events in their lives. Thus, it is impossible to determine whether memory of the traumatic event as a risk factor is specifically related to the trauma or, alternatively, is part of a more generalized pretraumatic vulnerability factor that relates to one’s autobiographical memory in general. Third, the instrument used for the evaluation of memory of the traumatic event does not cover the entire spectrum of the traumatic memory. Thus, participants may reexperience aspects of the trauma that are not captured by the instrument. Finally, it has often been suggested that PTSD patients are oversensitive to the adversity of the trauma and tend to overstate its etiological role in their psychopathology (7). Therefore, it could be argued that the retrospective appraisal of memory of the traumatic event might also be exaggerated. However, our findings indicate increased PTSD symptoms as early as 1 week after the trauma in respondents with memory of the traumatic event as well as the stability of memory of the traumatic event over time. Taken together, this makes it less likely that the findings of the current study are an artifact of retrospective exaggeration. Direct assessment with control for the degree to which it is important for the patient to remember the traumatic event may assist in clarifying this issue in future studies.

In conclusion, the results of our study indicate that in subjects with traumatic brain injury, memory of the traumatic event seems to be an important risk factor and predictor for subsequent PTSD. Future studies are needed to determine whether these findings are applicable to other traumatic conditions and whether interruption of the consolidation and the extinction of traumatic memories may be therapeutically beneficial.

References

7. van der Kolk B: Trauma and memory in traumatic stress, in The Effects of Overwhelming Experience on Mind, Body, and...
29. LeDoux JE: Fear and the brain: where have we been, and where are we going? Biol Psychiatry 1998; 44:1229–1238
The Impact of Comorbid Posttraumatic Stress Disorder on Short-Term Clinical Outcome in Hospitalized Patients With Depression

Paul E. Holtzheimer III, M.D.
Joan Russo, Ph.D.
Douglas Zatzick, M.D.
Christopher Bundy, M.D., M.P.H.
Peter P. Roy-Byrne, M.D.

Objective: Posttraumatic stress disorder (PTSD) is often comorbid with other psychiatric disorders but often goes unrecognized. The effects of PTSD comorbidity are unclear, especially in patients with severe mental illness. The authors assessed short-term clinical outcome in severely depressed psychiatric inpatients with and without comorbid PTSD.

Method: From patients hospitalized between 1995 and 2000, all patients with depression and comorbid PTSD (N = 587) were selected and matched with depressed patients without PTSD (N = 587). Clinical outcome was assessed with a semistructured, physician-administered battery. Differences between the two groups were examined, with overall burden of psychiatric illness entered as a covariate in the analyses.

Results: Relative to depressed patients without PTSD, depressed patients with PTSD had, at discharge, greater psychiatric symptom severity and higher levels of depression and hostility. Depressed patients with comorbid PTSD also had a significantly higher rate of being discharged against medical advice (odds ratio = 6.10, 95% CI 2.96–12.57).

Conclusions: PTSD comorbidity correlates with poorer short-term clinical outcome and greater likelihood of discharge against medical advice in severely depressed psychiatric inpatients. Better recognition of PTSD comorbidity may improve overall care of these patients.

Posttraumatic stress disorder (PTSD) is often present in patients with other psychiatric disorders, especially major depression (1–8). The National Comorbidity Survey found that 15.2% of subjects with major depressive disorder within the preceding 12 months had comorbid PTSD (2). In a sample of Bosnian refugees, 53% of those with major depressive disorder had PTSD (5). However, despite its frequency of occurrence, comorbid PTSD is often unrecognized in psychiatric patients, especially those with more severe illnesses, such as severe depression, bipolar disorder, and psychotic disorders (9–18).

The clinical relevance of PTSD comorbidity in depressed patients is largely unknown. Several studies have suggested that PTSD is associated with significant disability, impairment, and cost (4, 19–29). When PTSD is comorbid with another psychiatric illness, these effects may be even greater (2, 5, 6, 13, 25, 28, 30–33). However, few studies have specifically compared psychiatric patients with and without PTSD, and these studies have been limited to relatively small sample sizes. Also, since most studies have not controlled for total number of psychiatric diagnoses, it is possible that correlates of PTSD comorbidity merely reflect nonspecific correlates of increased psychiatric burden. Finally, the vast majority of studies to date have focused on general community samples, veterans, or refugee populations; few have systematically assessed the specific effects of PTSD comorbidity in a severely mentally ill population, where PTSD comorbidity may be more prevalent but underdiagnosed (9).

To better understand the impact of PTSD comorbidity in adult patients with severe mental illness, we performed an extensive, retrospective chart review of severely depressed psychiatric inpatients with and without comorbid PTSD. It is important to note that we were most interested in the presence of clinical comorbidity (as opposed to epidemiological comorbidity; see Kraemer [34] for a review); i.e., we intended to compare short-term clinical outcome (i.e., change in clinical status during psychiatric hospitalization) in depressed patients with and without PTSD at time of admission. The purpose of these exploratory analyses was to determine what differences, if any, in short-term clinical outcome were associated with PTSD comorbidity. It was presumed that any identified differences would help clarify the clinical importance of recognizing comorbid PTSD in severely depressed patients. Recognizing the limitations of a retrospective, case-control design, we carefully matched subjects in each group by age, gender, race/ethnicity, year of hospitalization, primary depressive diagnosis, and presence of substance use disorder diagnosis at time of admission. To better assess the specific associations of PTSD (as opposed to overall increased psychiatric burden), we used total number of psychiatric diagnoses as a covariate in our
analyses. We hypothesized that the short-term clinical outcome of depressed patients with PTSD would differ from that of depressed patients without PTSD.

**Method**

**Sample**

The initial study sample included all patients psychiatrically hospitalized between 1995 and 2000 at Harborview Medical Center in Seattle, Wash. All patients with a diagnosis of depression (including major depressive disorder, bipolar disorder [depressed phase], dysthymic disorder, depression not otherwise specified, or adjustment disorder with depressed mood) were first selected. All patients with a diagnosis of PTSD were then identified. Depressed patients with PTSD were matched with depressed patients without PTSD, selected from the original sample of depressed patients, according to age, gender, race/ethnicity (white versus nonwhite), year of admission (1995–1996, 1997–1998, 1999–2000), primary depression diagnosis (major depressive disorder, bipolar disorder [depressed phase], or other [dysthymic disorder, depression not otherwise specified, or adjustment disorder with depressed mood]), and current substance abuse or dependence (i.e., present at time of admission). Patients admitted on a planned basis as part of a structured program for patients with extremely high utilization of inpatient services were excluded. The study procedure was approved by the Institutional Review Board of the University of Washington Medical School and Harborview Medical Center.

**Data Collection/Assessments**

For all patients, a board-certified University of Washington faculty psychiatrist performed a standardized assessment battery within 24 hours of admission and discharge. This battery (35) included a demographic questionnaire, a semistructured psychiatric interview, the 23-item behaviorally anchored Psychiatric Symptom Assessment Scale (36), which includes domains similar to the Brief Psychiatric Rating Scale, and the DSM-IV Global Assessment of Functioning (GAF) scale. The semistructured diagnostic interview identified axis I and select axis II disorders (borderline personality disorder and antisocial personality disorder) using a symptom checklist based on DSM-IV criteria. This assessment battery has been shown to be reliable and valid for use with psychiatric inpatients in our setting (35). An extensive chart review was performed to confirm the presence of diagnostic criteria and the absence of any exclusionary criteria for the major diagnoses considered in this study (depression and PTSD). An automated hospital database was used to obtain length of stay and discharge status (regular or against medical advice).

Individual items on the Psychiatric Symptom Assessment Scale ranged from 0 (no severity) to 6 (most severe); the total score could range from 0 to 138. Cronbach's alpha for internal consistency reliability of the total Psychiatric Symptom Assessment Scale was 0.74 at admission and 0.78 at discharge. To assess depression severity, the scores of six Psychiatric Symptom Assessment Scale items associated with depression (withdrawal, motor retardation, helplessness/hopelessness, suicidality, depressed mood, blunted affect) were combined into a single score. To account for potential missing values on some items, the mean of the summed scores was calculated. In a similar fashion, an anxiety subscore was calculated by averaging the anxious mood, tension, excitement, and hyperactivity items into a single score, and a hostility subscore was calculated by averaging the uncooperativeness and hostility/aggression items into a single score. Factor analyses confirmed these subscales. Items in the depression subscore had an internal consistency reliability of 0.77 at admission and 0.78 at discharge. Items in the anxiety subscore had an internal consistency reliability of 0.73 at admission and 0.69 at discharge. Items in the hostility subscale had an internal consistency reliability of 0.73 at admission and 0.79 at discharge.

The assessment battery used in this study was not designed to establish the severity or time of onset of each psychiatric disorder diagnosed, nor was the source of PTSD clearly documented in each patient. However, to address the important question of how chronicity and type of PTSD might affect these results, a more detailed chart review was performed on a random sample of 178 (30%) of patients with depression and PTSD. In this subset, nine patients (5.1%) had acute PTSD that co-occurred with, or was preceded by, depressive symptoms; thus, approximately 95% of patients with PTSD in this study had chronic PTSD that preceded the current depression. Given the small proportion of patients with acute PTSD, we did not extend this chart review nor did we perform separate analyses on patients with acute PTSD. Additionally, among the 37 veterans with PTSD (6% of the PTSD sample), only 18 (3% of the PTSD sample) had combat-related PTSD. Given this small proportion, we did not compare results between PTSD subjects with or without combat trauma.

**Statistical Analyses**

For demographic comparisons, chi-square tests with a continuity correction were performed for categorical variables and independent sample t tests were used for age. For clinical severity analyses, the two groups were compared at admission and discharge on total Psychiatric Symptom Assessment Scale score; depression, anxiety, and hostility subscale scores; and GAF score. Logistic regression was used for categorical variables, and analysis of covariance was used for continuous variables. Number of psychiatric diagnoses was used as a covariate for all analyses. Substance abuse/dependence was not considered a primary or comorbid psychiatric diagnosis, since this was used as a primary matching variable. Admission values were used as covariates in analyses of discharge variables.

Characteristics of the hospital stay were analyzed by comparing the two groups on likelihood of involuntary hospitalization, likelihood of discharge against medical advice, and length of stay. Length of stay analyses used number of diagnoses and admission clinical severity (total Psychiatric Symptom Assessment Scale score) as covariates. Separate comparisons were made for patients admitted voluntarily and involuntarily. Also, patients discharged against medical advice were excluded from these analyses to decrease bias from unexpectedly short hospital stays.

All statistical analyses were two-tailed. Given the exploratory nature of these analyses, a significance level of p=0.01, without correction for multiple comparisons, was used.

**Results**

A total of 4,182 psychiatrically hospitalized patients with depression were initially identified. Within this group, 587 patients (14%) with comorbid PTSD were matched with 587 depressed patients without PTSD to provide the study sample. Patients were matched in terms of gender, race/ethnicity, age, year of admission, depression diagnosis, and current substance abuse/dependence. There were no significant differences between depressed patients with or without PTSD on demographic variables (Table 1), although there was a tendency for fewer patients with PTSD to be employed.

On admission, there were no differences between the two groups on clinical severity (Table 2). However, at discharge the PTSD group had a significantly higher total score on the Psychiatric Symptom Assessment Scale and
higher scores on the depression and hostility subscales than did patients without PTSD (Table 3). There also tended to be lower discharge GAF scores in the PTSD group. Because differences in length of stay and rates of discharge against medical advice could account for differences in discharge clinical severity, we performed post hoc analyses covarying for these variables. Covarying for length of stay did not alter the statistical significance of any of the original findings. Among patients discharged regularly (i.e., not against medical advice), patients with PTSD continued to show higher total scores on the Psychiatric Symptom Assessment Scale ($F=9.09$, $df=1$, 982, $p=0.003$) and higher depression subscale scores ($F=8.52$, $df=1$, 961, $p=0.004$) than patients without PTSD. The hostility subscale score remained higher, but not significantly so ($F=2.96$, $df=1$, 985, $p<0.09$). There continued to be no significant differences in anxiety subscale scores or discharge GAF scores. Given the small proportion of patients discharged against medical advice (<7% of total sample), we did not perform post hoc analyses in this group.

Length of hospitalization and the proportion of patients hospitalized involuntarily did not differ significantly between the groups (Table 4). However, likelihood of discharge against medical advice was significantly higher in the PTSD group, with an odds ratio of 6.10 (95% CI=2.96–12.57). When separate analyses for likelihood of discharge against medical advice were done for voluntary ($N=67$ [with PTSD: $N=58$, no PTSD: $N=9$]) and involuntary ($N=11$ [with PTSD: $N=8$, no PTSD: $N=3$]) admission patients, this difference remained significant for voluntary admission patients (odds ratio=7.26, 95% CI=3.22–16.35; Wald’s $\chi^2=22.92$, $df=1$, $p<0.0001$) but not involuntary admission patients (odds ratio=2.38, 95% CI=0.44–12.95; Wald’s $\chi^2=1.01$, $df=1$, $p<0.32$). However, the small sample size of involuntary patients discharged against medical advice limits the power of this comparison.

While our assessment battery was not designed to obtain data on the chronicity or source of PTSD, these features were investigated in a subset of the depressed patients with PTSD. Of 178 patients (30% of PTSD sample), nine (5.1%) had acute PTSD, 53 (29.8%) had chronic PTSD related to childhood trauma, 48 (27.0%) had chronic PTSD related to multiple traumas (often including childhood trauma), and 39 (21.9%) had chronic PTSD related to adult trauma. For 29 patients with chronic PTSD (16.3%), the source of the PTSD could not be identified from the chart review.

**Discussion**

To date, this is one of the largest studies of PTSD comorbidity in depressed patients and one of the first to systematically investigate the effects of PTSD comorbidity on clinical outcome in a severely depressed, hospitalized population. In this study, depressed inpatients with comorbid PTSD were more psychiatrically impaired, more depressed, and more hostile at time of discharge than depressed patients without PTSD. Also, hospitalized patients with depression and PTSD were much more likely to be discharged against medical advice.
These results are consistent with previous reports that have suggested that psychiatric patients with comorbid PTSD are more symptomatic and more impaired than individuals without PTSD comorbidity (2, 5, 6, 13, 19, 20, 25, 28, 30–33). It is important to note that in this study, such effects were noted only at discharge (versus admission), suggesting that hospitalized, depressed patients with PTSD have poorer short-term clinical outcomes than patients without PTSD. Whether these effects are specific to PTSD or related to overall psychiatric impairment has been questioned (37). Since overall burden of psychiatric illness was entered into analyses as a covariate, our results indicate that PTSD comorbidity is independently associated with poorer clinical outcome in hospitalized, depressed patients.

One of the most striking findings in this study is the greatly increased likelihood of discharge against medical advice among patients with comorbid PTSD. Depressed patients with PTSD were more than six times as likely to leave the hospital against medical advice than those without PTSD. This finding was present even in the absence of differences between the two groups on a number of variables known to predict discharge against medical advice (age, gender, race, substance abuse/dependence, severity of medical comorbidity) (38). While a previous study found that primary psychiatric diagnosis did not predict discharge against medical advice (38), our results indicate that PTSD comorbidity may have specific and profound effects on the likelihood of discharge against medical advice in depressed patients. This finding adds to growing evidence that patients with PTSD may be particularly difficult to engage in mental health treatment (19, 39–42) and has important implications for treatment planning for psychiatric patients with comorbid PTSD.

In this study, hospitalized depressed patients with PTSD did not differ significantly from depressed patients without PTSD on a number of demographic variables, nor was the presence of PTSD comorbidity associated with differences in clinical severity on admission. In particular, we did not replicate findings reported by other groups that PTSD comorbidity is associated with increased suicidality (2, 13, 25, 30). However, the similarity in clinical presentation at admission, including history of suicide attempts, between hospitalized depressed patients with and without PTSD likely represents a threshold effect for admission in a population of severely ill patients already at high risk for suicide.

It is interesting that depressed patients with comorbid PTSD did not demonstrate higher anxiety at admission or discharge than did depressed patients without PTSD. The absence of a difference in anxiety in this sample is somewhat surprising, given that patients were stratified by presence or absence of a comorbid anxiety disorder (i.e., PTSD). A possible explanation of this finding is that patients with PTSD in this study (most of whom had very chronic PTSD that preceded the diagnosis of depression) overall were more likely to demonstrate withdrawal and avoidance (perhaps demonstrated by the higher rate of discharge against medical advice) and less likely to show obviously increased anxiety (as measured by the anxious mood, tension, excitement, and hyperactivity items of the Psychiatric Symptom Assessment Scale).

Taken together, these results strongly suggest that for severely depressed patients, comorbid PTSD is associated

### TABLE 2. Clinical Severity at Admission for Severely Depressed Psychiatric Inpatients With or Without Comorbid PTSD

<table>
<thead>
<tr>
<th>Measure</th>
<th>PTSD</th>
<th>No PTSD</th>
<th>Analysis&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Valid Nb&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Score</td>
<td>Mean</td>
</tr>
<tr>
<td>Psychiatric Symptom Assessment Scale Total</td>
<td>566</td>
<td>33.33</td>
<td>11.48</td>
</tr>
<tr>
<td>Depression</td>
<td>553</td>
<td>2.96</td>
<td>1.02</td>
</tr>
<tr>
<td>Anxiety</td>
<td>559</td>
<td>1.36</td>
<td>1.02</td>
</tr>
<tr>
<td>Hostility</td>
<td>565</td>
<td>0.85</td>
<td>1.23</td>
</tr>
<tr>
<td>Global Assessment of Functioning Scale</td>
<td>572</td>
<td>27.17</td>
<td>12.29</td>
</tr>
</tbody>
</table>

<sup>a</sup> Analysis of covariance, with number of psychiatric diagnoses used as a covariate.

<sup>b</sup> Number of subjects with data for variable in chart.

### TABLE 3. Clinical Severity at Discharge for Severely Depressed Psychiatric Inpatients With or Without Comorbid PTSD

<table>
<thead>
<tr>
<th>Measure</th>
<th>PTSD</th>
<th>No PTSD</th>
<th>Analysis&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Valid Nb&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Score</td>
<td>Mean</td>
</tr>
<tr>
<td>Psychiatric Symptom Assessment Scale Total</td>
<td>507</td>
<td>13.16</td>
<td>8.72</td>
</tr>
<tr>
<td>Depression</td>
<td>498</td>
<td>0.86</td>
<td>0.67</td>
</tr>
<tr>
<td>Anxiety</td>
<td>500</td>
<td>0.75</td>
<td>0.69</td>
</tr>
<tr>
<td>Hostility</td>
<td>512</td>
<td>0.57</td>
<td>1.09</td>
</tr>
<tr>
<td>Global Assessment of Functioning Scale</td>
<td>473</td>
<td>48.45</td>
<td>11.21</td>
</tr>
</tbody>
</table>

<sup>a</sup> Analysis of covariance, with number of psychiatric diagnoses and admission values used as covariates in all analyses.

<sup>b</sup> Number of subjects with data for variable in chart.
TABLE 4. Inpatient Hospitalization Characteristics for Severely Depressed Psychiatric Inpatients With or Without Comorbid PTSD

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PTSD</th>
<th>No PTSD</th>
<th>Analysis a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Valid N 0</td>
<td>Valid N 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>With Characteristic</td>
<td>With Characteristic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Involuntary admission</td>
<td>587</td>
<td>56</td>
<td>9.5</td>
</tr>
<tr>
<td>Discharged against medical advice</td>
<td>587</td>
<td>66</td>
<td>11.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of stay (days)c</td>
<td>All patients</td>
<td>521</td>
<td>10.8</td>
<td>6.8</td>
<td>575</td>
<td>9.9</td>
<td>6.0</td>
</tr>
<tr>
<td>Voluntary</td>
<td>48</td>
<td>12.6</td>
<td>8.6</td>
<td>78</td>
<td>16.4</td>
<td>9.7</td>
<td>1.80</td>
</tr>
<tr>
<td>Involuntary</td>
<td>473</td>
<td>9.6</td>
<td>5.6</td>
<td>497</td>
<td>9.9</td>
<td>5.8</td>
<td>0.42</td>
</tr>
</tbody>
</table>

a Logistic regression used for categorical variables; analysis of covariance used for continuous variables. Number of psychiatric diagnoses used as a covariate in all analyses.

b Number of subjects with data for variable in chart.

c Patients discharged against medical advice excluded from length of stay analyses; baseline severity (total Psychiatric Symptom Assessment Scale score) used as a covariate.

with a poorer outcome during psychiatric hospitalization, including a greater likelihood for discharge against medical advice. This emphasizes the importance of recognizing PTSD comorbidity in such patients and raises the question of whether depressed patients with comorbid PTSD might benefit from more specialized treatment during psychiatric hospitalization.

Certain limitations must be considered when interpreting these results. Most important, diagnoses were not obtained through the use of a fully structured diagnostic instrument such as the Structured Clinical Interview for DSM-IV Axis I Disorders (43). However, our goal in this study was to include data from a large number of severely ill, hospitalized patients in order to generate a sample size that would allow for analyses not possible in previous smaller studies. To do this, we included naturalistic data collected on all patients hospitalized over a 6-year period, using the assessment battery described earlier. The semi-structured diagnostic interview used in this battery has been shown to be valid in other settings and reliable for use with psychiatric inpatients in our setting (35, 44). In addition, an extensive chart review was used to confirm presence or absence of PTSD and depression diagnoses in the study sample. While the likelihood of false positives is potentially higher with this approach, the accuracy of diagnoses in this study is likely to be greater given the high prevalence of mood disorders and PTSD in an inpatient setting. Since PTSD is commonly underdiagnosed, it is certainly possible that some depressed patients may have been incorrectly categorized as not having comorbid PTSD.

With this data set, we were not able to control for type and number of traumatic experiences or for duration of PTSD (although a subsample analysis suggested about 95% of patients had chronic PTSD that preceded the current depression). The severity of PTSD was also not consistently tracked, so the effects of PTSD severity on the outcome measures in this study are unknown. Also, the nature of the data collection did not allow for meaningful analysis of a number of health care utilization measures of interest (e.g., number of hospitalizations or frequency of emergency room visits), since a bias toward patients admitted earlier would have been present.

Retrospective chart reviews may be limited by data missing because of inadequate documentation. Missing data were encountered in this study and account for the variability in sample size among the different comparisons. It is important to note that there were no significant differences in the amount of missing data between the two groups for these analyses. Also, it should again be noted that this retrospective chart review allowed for the collection of an extensive dataset on a very large sample of patients.

Given the nature of the data used in this study, a case/control design was employed for analyses. Case/control studies may be limited by lack of randomization resulting in systematic differences between groups. To limit this, we employed extensive matching of depressed patients with and without PTSD. Accordingly, demographic comparisons suggested the groups were quite similar on a number of variables that could have confounded the results (e.g., age, gender, substance abuse/dependence, primary depressive diagnosis, year of admission) (Table 1).

Despite these limitations, these results clearly show that comorbid PTSD is associated with important and independent differences in short-term clinical outcome, especially rate of discharge against medical advice, in patients with severe depression. These findings strongly suggest that identifying PTSD comorbidity is relevant to the care of psychiatric inpatients with depressive illness. A next step will be to investigate whether identification and treatment of PTSD in these patients can ameliorate the effects of PTSD comorbidity observed in this hospitalized cohort of severely depressed patients.
The authors thank Ms. Lynn DeJessa for her assistance with editing and preparing this manuscript.

References

37. Zlotnick C, Franklin CL, Zimmerman M: Is comorbidity of posttraumatic stress disorder and borderline personality disorder...
41. Schwarz ED, Kowalski JM: Malignant memories: reluctance to utilize mental health services after a disaster. J Nerv Ment Dis 1992; 180:767–772
Relief of Expressed Suicidal Intent by ECT: A Consortium for Research in ECT Study

Charles H. Kellner, M.D.
Max Fink, M.D.
Rebecca Knapp, Ph.D.
Georgios Petrides, M.D.
Mustafa Husain, M.D.
Teresa Rummans, M.D.
Martina Mueller, Ph.D.
Hilary Bernstein, L.I.S.W.
Keith Rasmussen, M.D.
Kevin O’Connor, M.D.

Glenn Smith, Ph.D.
A. John Rush, M.D.
Melanie Biggs, Ph.D.
Shawn McClintock, M.S.
Samuel Bailine, M.D.
Chitra Malur, M.D.

Objective: This study assessed the incidence, severity, and course of expressed suicidal intent in depressed patients who were treated with ECT. The data are from the first phase of an ongoing, collaborative multicenter study, the overall aim of which was to compare continuation ECT with pharmacotherapy in the prevention of relapse after a successful course of ECT.

Method: Suicidal intent, as expressed by patients during an interview, was scored at baseline and before each ECT session with item 3 on the 24-item Hamilton Depression Rating Scale in 444 patients with unipolar depression.

Results: One hundred thirty-one patients (29.5%) reported suicidal thoughts and acts (score of 3 or 4) at baseline. Scores decreased to 0 after 1 week (three ECT sessions) in 38.2% of the patients, after 2 weeks (six ECT sessions) in 61.1%, and in 80.9% at the end of the course of treatment.

Conclusions: Expressed suicidal intent in depressed patients was rapidly relieved with ECT. Evidence-based treatment algorithms for major depressive mood disorders should include dichotomization according to suicide risk, as assessed by interview. For patients at risk, ECT should be considered earlier than at its conventional “last resort” position.

Suicide and suicide attempts are risks of the major psychiatric illnesses, with mortality rates markedly higher than for the general population (1). A lifetime risk of 15% for suicide was the conclusion of a meta-analysis of hospitalized patients by Guze and Robins (2). A reanalysis of studies of suicide in affective illness in three groups of patients—outpatients, inpatients, and suicidal patients—found a risk of those ever hospitalized of 8.6% (3). Other studies confirmed the risk of suicide to be particularly high in hospitalized depressed patients (4–6). Profound hopelessness, hypochondriacal ruminations or delusions, and thoughts of suicide or self-harm during depression predict future suicide (7).

Antidepressant medications are the principal agents used to treat affective disorders today. Their impact on suicide risk is not well defined, however, and they are generally viewed as less effective than ECT in relieving depression and suicidal thoughts. Hordern et al. (8) reported no suicides at the 6-month follow-up in 34 women treated with ECT but found two in the 84 patients treated with antidepressants (2.4%). Avery and Winokur (9) assessed suicidal behavior in the 6 months following the treatment of depression in 519 patients. Suicide attempts were recorded in 0.8% of the ECT patients compared to 4.2% of those who had received “adequate” and 7% “inadequate” antidepressant medication treatment. Khan et al. (10) found antidepressant drugs to be no more effective than placebo in reducing suicide rates in seriously ill depressed patients.

The experience with ECT is also unclear. A comparison of the frequency of suicides in different decades found decreased rates when ECT was the dominant treatment for mental illness (11). An examination of the records of 1,397 completed suicides in Finland within a 12-month period showed only two patients having received ECT (12). In another report, Sharma (13) identified 45 psychiatric inpatients in a Canadian provincial hospital who committed suicide and compared their records with an equal number of patients matched for gender, age, and admission diagnosis (13). He reported no difference in ECT use in the two samples, finding no particular benefit for ECT.

Prudic and Sackeim (14) examined the changes in item 3 (expressed suicidal intent) of the Hamilton Depression Rating Scale in 148 patients referred for ECT. Both suicide and mortality rates were reduced with treatment. The overall average score on item 3 was 1.8 at baseline and was reduced to 0.1 in 72 responders and to 0.9 in 76 nonresponders. For the total sample, there was a greater de-
crease in the suicide item scores than the overall Hamilton depression scale scores.

National psychiatric associations in Great Britain and the United States (15, 16) and recent assessments by the Canadian Agence d’Évaluation des Technologies et des Modes d’Intervention en Santé (17) and the U.K. National Institute for Clinical Excellence (18) cited the reduction of suicide risk as a justification for the use of ECT.

We assessed suicide risk in an ongoing collaborative multicenter study comparing the efficacy of continuation ECT with continuation pharmacotherapy in unipolar depressed patients after response to a course of ECT.

Method

Study Design

The Consortium for Research in ECT Continuation ECT Study is a multisite, collaborative study funded by the National Institute of Mental Health to compare the efficacy of continuation ECT and continuation pharmacotherapy (lithium plus nortriptyline). Patients with unipolar major depression were referred for ECT at four institutions to receive an acute course of thrice-weekly bilateral ECT sessions (acute phase) (19). ECT is given with a standardized protocol with a Thymatron DGx device (Somatics, Inc., Lake Bluff, Ill.). The patients whose condition remits and remains remitted for 1 week without treatment are randomly assigned to receive either continuation ECT or continuation pharmacotherapy (nortriptyline plus lithium) (continuation phase). In the 6-month continuation phase, the patients receive bilateral ECT weekly for 4 weeks, biweekly for 8 weeks, and monthly for 2 months.

Depressive symptoms are assessed at baseline and three times weekly during the course of treatment (24–36 hours after each ECT session) with the 24-item Hamilton depression scale (20). Remission in the acute phase is defined by two consecutive Hamilton depression scale ratings ≤10 and a ≥60% reduction in Hamilton depression scale total score. The reliability of these ratings was checked by a quality assurance program in which a doctoral-level psychologist reviewed and co-rated videotapes of interviews from each site. All raters participated in a co-rating of interviews at least three times a year. When ratings were not within the acceptable range, the rater participated in customized remediation sessions.

There is no prescribed minimum or maximum number of ECT sessions; however, to be declared a nonremitter, a patient must have received at least 10 ECT treatments and reached a plateau relative to the decline of Hamilton depression scale total scores, indicating no further clinical improvement. Patients who withdraw consent or for whom ECT is discontinued for clinical or other reasons before receiving 10 ECT sessions are considered dropouts.

Study Group

Patients with DSM-IV defined major depressive disorder (major depressive disorder as assessed by the Structured Clinical Interview for DSM-IV Axis I Disorders criteria [21]), with age ranges of 18–85 years, who were referred for ECT and who had a minimum baseline Hamilton depression scale total scores of 21 were eligible for study inclusion. Exclusion criteria included schizophrenia, bipolar disorder, neurological illness, or illicit substance dependence within 1 year and ECT within 3 months. To date, 444 patients with unipolar major depression have entered the study, with 355 (80%) completing the acute course of ECT and 89 dropping out (20%). Reasons for dropping out during the acute phase included patient withdrawal of consent (N=37), adverse events (N=33), and protocol violations (N=19). Reasons for patient-initiated withdrawal (N=37) included concerns about cognitive effects (N=5), no perceived treatment effects (N=1), the patient perceiving that no further treatments were needed (N=5), preferred medications (N=6), wanting to go home (N=2), and no reason given (18). Adverse events were confusion or memory problems (N=20), intercurrent medical condition (N=9), and other psychiatric illnesses (N=4). The study group was 31.8% male, 91.7% Caucasian, 29.7% with psychotic depression. The mean Hamilton depression scale score at baseline was 35.1 (SD=7.2). Demographic characteristics of the study group are representative of patients typically referred for ECT.

Acute-Phase ECT

A stimulus dose titration procedure was performed at the first treatment to determine seizure threshold, and subsequent stimulus dosing was at 1.5 times the seizure threshold. The patients were treated with a standard brief-pulse ECT device with bitemporal electrode placement.

Outcome

This report considers the change in expressed suicidal intent, as recorded in item 3 of the Hamilton depression scale. This item was scored 0=absent; 1=feels life is empty or is not worth living; 2=recurrent thoughts or wishes of death of self; 3=active suicidal thoughts, threats, gestures; and 4=serious suicide attempt. Scores of 3 and 4 were evidence of expressed suicide intent (0–2 [low suicide rating group] and 3–4 [high suicide rating group]). The primary outcome was resolution of the suicide score to 0 (score resolved to 0/did not resolve to 0). For primary analyses, those in the high suicide rating group were followed for occurrence of the primary outcome (item 3 score resolved to 0).

Statistical Analysis

Standard descriptive measures (means and standard deviations for continuous variables, proportions for categorical variables) were used to describe baseline demographic and clinical characteristics of the total study group and the high versus low suicide rating groups. Baseline characteristics were compared across the low/high suicide rating groups with either a two-sample t test (continuous variables) or a chi-square test (categorical variables). For the primary analysis, a frequency distribution describing proportion of patients in the high rating group whose suicidality scores resolved (became 0) after successive ECT treatments was determined. In additional secondary analyses, percent change in item 3 means and corresponding Hamilton depression scale total score means for the total study group after 1 week (three ECT sessions) and 2 weeks (six ECT sessions) of treatment were also determined. The statistical significance of the decline in item 3 values over the course of the first three ECT sessions (1 week of treatment) and over the course of the first six ECT sessions (within 2 weeks of treatment) was assessed with mixed-models analysis (SAS PROC MIXED with unstructured covariance [SAS Institute, Cary, N.C.]), which evaluates the longitudinal profile (e.g., slope) in repeated correlated measures over time. Characteristics of the groups whose expressed suicidal intent scores resolved (became 0) and whose scores did not resolve were described with standard descriptive methods and were compared with either a two-sample t test (continuous variables) or a chi-square test (categorical variables). In exploratory analyses, we investigated the relationship of psychosis status and age to resolution of expressed suicidal intent with logistic regression with expressed intent resolved or not resolved as the dependent variable. Psychosis status and age were considered both separately (univariate analyses) and together (multivariate analyses). Odds ratios and corresponding 95% confidence intervals (CIs) are reported from these analyses. No correction for multiple analyses.
was used; therefore, secondary and exploratory results must be evaluated cautiously in terms of type I error rate.

Results

The remission rate for depression in the 355 patients who completed the course of treatment was 85.6% (N=304). For the total study group (N=444) (treating dropouts as nonremitters), the remission rate was 68.5% (N=304). The average Hamilton depression scale score at baseline was 35.1 (SD=7.2) for the full group, with an average change from baseline of 24.8 (SD=10.2) for the full group and 26.9 (SD=9.1) for those who completed the acute course of ECT.

On item 3 of the Hamilton depression scale the 444 patients referred for ECT, 118 (26.6%) received a score of 3 for having active suicidal thoughts, actions, or gestures, and 13 (2.9%) received a score of 4 for reporting a suicidal event during the current episode. The baseline means for the item 3 ratings were 1.70 (SD=1.16) for the full group, 1.64 (SD=1.18) for the 355 patients who completed the full treatment course, and 1.92 (SD=1.00) for the 89 dropouts.

Patients in the high expressed suicidal intent group were younger, on average, than those in the low group (mean=47.2, SD=13.8, versus mean=59.1, SD=16.6) (p<0.001) (Table 1). Although there were fewer psychotic patients and women in the high expressed suicidal intent group, the differences were not significant (p=0.37 and p=0.09, respectively) (Table 1). There were no differences in the dropout rates for those in the high compared to the low expressed suicidal intent groups (22% versus 19%) (χ²=0.51, df=1, p=0.48).

Of the 131 patients in the high expressed suicidal intent group, the rating of 106 patients (80.9%) ultimately dropped to 0. This occurred in 15.3% (N=20) after one ECT session; in 38.2% (N=50) after three ECT sessions (1 week); in 61.1% (N=80) after six ECT sessions (2 weeks); and in 76.3% (N=100) after nine ECT sessions (3 weeks) (Table 2). Among 102 patients in the high expressed suicidal intent group who completed the acute course of ECT, 87.3% (N=89) had their scores drop to 0 by the end of the treatment course, with approximately 63.7% reaching resolution after six ECT sessions (within 2 weeks). For the 13 patients with a rating of 4 (suicide attempt) at baseline, all saw their ratings drop to 0 by the end of their treatment.

Among the 25 patients in the high expressed suicidal intent group whose ratings did not resolve to 0, 48% (N=12) dropped out before receiving an adequate course of treatment and, of the remaining 13 patients, 46% (N=6) had a rating of 1 at the end of the acute course. Over half (N=7) of the 13 patients rated 0 during the treatment course but did not retain it at the final rating. The patients whose expressed suicide intent ratings did not resolve were younger than those whose ratings resolved (t=−2.3, df=129, p<0.02) but were not different with respect to percent with psychosis or baseline severity of illness (Table 3).

After six ECT sessions (2 weeks), 66% (N=35) of 53 older patients (≥50 years) had expressed suicidal intent scores of 0 compared to 58% (N=45) of 78 younger (<50 years) patients; 89% (N=47) of the older group eventually reached ratings of 0 compared to 76% (N=59) of the younger group. The odds of having the expressed suicide intent rating resolve to 0 was 2.5 times higher for the older compared to the younger group (odds ratio=2.5, 95% confidence interval=0.9–6.8). When we treated age as a continuous variable in the logistic regression model, there was a significant relationship between the outcome (expressed suicidal intent rating resolved or did not resolve) and age (p<0.03).

At baseline, the correlation between item 3 ratings and Hamilton depression scale total score was −0.01 (Pearson correlation, p=0.87) among the group in the high rating category. The correlation between change from baseline for the item 3 rating and Hamilton depression scale total score was 0.45 (Pearson correlation, p<0.001) for this

---

**TABLE 1. Demographic and Baseline Clinical Characteristics of Depressed Subjects With Expressed Suicidal Intent**

<table>
<thead>
<tr>
<th>Baseline Variable</th>
<th>Total Study Group (N=444)</th>
<th>Subjects With Expressed Suicidal Intent</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>Low (N=313) %</td>
<td>High (N=131) %</td>
</tr>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>t (df=442)</td>
</tr>
<tr>
<td>Female gender</td>
<td>303 (68.2)</td>
<td>221 (70.6)</td>
<td>82 (62.6) 2.7 0.09</td>
</tr>
<tr>
<td>Psychotic</td>
<td>132 (29.7)</td>
<td>97 (30.1)</td>
<td>35 (26.7) 0.8 0.37</td>
</tr>
<tr>
<td>Caucasian</td>
<td>407 (91.7)</td>
<td>287 (91.7)</td>
<td>120 (91.6) 0.0 0.98</td>
</tr>
<tr>
<td>Completed acute course</td>
<td>355 (80.0)</td>
<td>253 (80.8)</td>
<td>102 (77.8) 0.5 0.48</td>
</tr>
<tr>
<td>Age (years)</td>
<td>55.6 (16.8)</td>
<td>59.1 (16.6)</td>
<td>47.2 (13.8) 7.2 &lt;0.001</td>
</tr>
<tr>
<td>Number of ECT sessions</td>
<td>7.3 (3.4)</td>
<td>7.2 (3.4)</td>
<td>7.5 (3.2) −0.7 0.47</td>
</tr>
<tr>
<td>24-item Hamilton depression rating scale score</td>
<td>35.1 (7.2)</td>
<td>33.6 (6.6)</td>
<td>38.6 (7.3) −7.1 &lt;0.001</td>
</tr>
</tbody>
</table>

a Patients with low levels of expressed suicidal intent had scores of 1 or 2, and patients with high levels had scores of 3 or 4 on item 3 of the 24-item Hamilton Depression Rating Scale.

b Comparing proportions (%) for lower versus higher expressed suicidal intent groups.

c Independent sample (pooled test) comparing means for lower versus higher expressed suicidal intent groups.

---

group. Within the first week of treatment (three ECT sessions), the average percent reduction from baseline in item 3 scores for the total study group was 56% compared to a relative change of 52% for the Hamilton depression scale total score. After 2 weeks of treatment (six ECT sessions), the expressed suicide intent rating item had a relative reduction from baseline of 78% (75% for Hamilton depression scale total score).

Two patients died by suicide in the course of the study, one each at different sites. Both were white men, 76 and 80 years of age, and were diagnosed with a major depressive disorder without psychosis. At the outset, one patient had a Hamilton depression scale of 29 with an expressed intent that was scored 1 on item 3. He received seven ECT sessions and ended the course with a Hamilton depression scale score of 8 and a score of 0 on item 3. He was in the 1-week interim phase awaiting random assignment, and on day 6 after the last ECT session, he died by gunshot wound. The second patient had a baseline Hamilton depression scale score of 22 and expressed no suicidal intent, scoring 0 on item 3. After 10 ECT sessions, his Hamilton depression scale score was 8, and 1 week later, he was randomly assigned to receive continuation ECT. Two weeks later, after the second continuation ECT session (Hamilton depression scale score of 16 and item 3 score of 2), he successfully overdosed with medications.

Discussion

These data, from a large group of severely depressed patients, most of whom were hospitalized, document a rapid reduction in expressed suicidal intent in patients treated with ECT. The patients were referred at four geographically dispersed academic medical centers and are typical of depressed patients who receive ECT after extensive medication and psychotherapy trials. Expressed suicidal intent, as reflected in scores of 3 and 4 on item 3 of the Hamilton depression rating scale, was present in 29.5% of the treatment group of 444. This incidence of expressed suicidal intent is considerably larger than in the usual populations treated with medications (22).

Two of 444 patients died by suicide, both in the follow-up period after the index ECT course and both on outpatient status. We find the benefits of ECT to be immediate, but they may not persist unless treatment is sustained, a conclusion also reached by others (11, 14). More intensive and extensive treatment is needed to consolidate the benefits of antidepressant treatment.

Suicide risk is high after discharge from hospital care for depression. In a study of depressed patients at 3 months, 1 year, and 2 years after admission, Oquendo et al. (23) reported that the persistence of major depression in the follow-up period increased the risk of a suicide attempt sevenfold. Antidepressant treatment in the follow-up period was largely inadequate, and patients with a prior suicide attempt did not receive more vigorous care than nonattempters. For each suicide attempt in the subject’s history, the risk for an attempt in the follow-up period increased by 30%.

Two cohort studies also reported higher suicide rates in depressed patients immediately after hospital discharge (24, 25). A third study did not find an association (26). The author noted “that no specific diagnosis seems to stand out suggest[ing] that an improvement in the quality of care for all discharged patients is necessary to reduce suicide rates.” We would add that effective remission of mood disorder and psychosis should reduce the incidence of suicide after hospital discharge.

These observations are relevant to the development of treatment algorithms for major depressive illness. Our present treatment algorithms do not reflect the efficacy of ECT (27). They offer treatments of lesser, and even some unproven, efficacy before recommending ECT, which is usually cited as the “last resort.” Such approaches are justified by the controversial nature of ECT, the reluctance of patients to accept it without having tried all other treatment courses, and the unavailability of ECT at major treatment centers. Indeed, in some states, notably California and Texas, the last resort option is codified into law (28, 29). Whatever the objections to the use of ECT may be, the present data on the rapid resolution of expressed suicidal intent warrant greater attention in evidence-based treatment algorithms.

Just as it has been advantageous to dichotomize depressive patients as to the presence of psychosis, offering different treatment options for the two groups, it may be advantageous to dichotomize patients at risk of suicide. For those at high suicide risk or with recent suicide attempts, especially those who appear so suicidal as to require continuing observation, ECT should be considered early in the algorithm.

The presence of psychosis in depression is an additional risk factor for suicide. Roose et al. (4) retrospectively studied 22 suicides in patients with major affective disorders. Their 14 patients with delusional unipolar depression

<table>
<thead>
<tr>
<th>Session Number</th>
<th>Total Study Group (N=131)</th>
<th>Completers (N=102)</th>
<th>Dropouts (N=29)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>15.3</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>39</td>
<td>29.8</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>38.2</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>48.9</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>72</td>
<td>55.0</td>
<td>58</td>
</tr>
<tr>
<td>6</td>
<td>80</td>
<td>61.1</td>
<td>65</td>
</tr>
<tr>
<td>7</td>
<td>86</td>
<td>65.6</td>
<td>71</td>
</tr>
<tr>
<td>8</td>
<td>90</td>
<td>68.7</td>
<td>75</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>76.3</td>
<td>84</td>
</tr>
<tr>
<td>10</td>
<td>101</td>
<td>77.1</td>
<td>85</td>
</tr>
<tr>
<td>&gt;10</td>
<td>106</td>
<td>80.9</td>
<td>89</td>
</tr>
<tr>
<td>No ECT</td>
<td>25</td>
<td>19.1</td>
<td>13</td>
</tr>
</tbody>
</table>
were 5.3 times as likely to complete suicide as their unipolar depressed patients without delusions. Patients with psychotic depression respond better and more rapidly to real rather than sham ECT (30, 31). They also respond more rapidly and better than nonpsychotic depressed patients to ECT (19, 32).

A limitation of the current study was the use of a single item to assess a behavior as complex as expressed suicidal intent. And, in the absence of a concurrent comparison group, we depend on historical comparisons with antidepressant medications.

Advantages of the current study include the large group, multiple sites, standardized single treatment, well-trained raters, rigorous nature of the assessments, and the limitations in variability of additional medications and rescue treatments, thereby defining a rigorous study group with homogeneous treatment characteristics. The use of bilateral electrode placement, the monitoring of the EEG of each treatment to ensure adequate treatments, and the use of standardized dosing schedules helped to improve treatment outcome.

More and more emphasis is being placed on outcomes in psychiatric practice. The most recent large-scale meta-analysis of the efficacy and safety of ECT in depressive disorders reports real ECT to be more effective than simulated ECT and more effective than pharmacotherapy (33). The authors conclude, “There is a reasonable evidence base for the use of ECT; it does not rest simply on anecdote, habit, and tradition….ECT remains an important treatment option for the management of severe depression.” Indeed, a recent report on the quality of health care delivered to adults in the United States (34) finds a disturbing failure of patients to get the recommended medical treatment for their illness almost half the time. The efficacy of ECT in treating depression is largely ignored in treatment algorithms, limiting the chance that severely ill, depressed patients will get an available effective treatment.

The profession often cites suicide risk as a justification for ECT (15, 17). The present study encourages earlier consideration of ECT in the course of the treatment of severely depressed patients than is presently offered in the professional literature.

Conclusions

Considering the risk of suicide and the delayed efficacy of medications in severely depressed patients, the present practice of recommending ECT as a last resort in expert treatment algorithms unnecessarily puts suicidal psychiatrically ill patients at substantial risk (23, 35). The common recommendation that medication be offered patients with severe depression, regardless of the symptom of expressed suicidal intent, is not justified. It is prudent to consider differentiating treatment options on a measure of expressed suicide intent.

From the University of Medicine and Dentistry of New Jersey–New Jersey Medical School; Zucker Hillside Hospital, North Shore–Long Island Jewish Health System, Glen Oaks, N.Y.; the Medical University of South Carolina, Charleston, S.C.; the University of Texas Southwestern Medical Center, Dallas; and Mayo Clinic and Foundation, Rochester, Minn. Address reprint requests to Dr. Kellner, UMDNJ–New Jersey Medical School, 183 South Orange Ave., F1436, Newark, NJ 07103; kellner@umdnj.edu (e-mail).

Supported by NIMH grants MH-55495 (to New Jersey Medical School), MH-55484 (to Mayo Clinic and Foundation), MH-55486 (to Long Island Jewish Health System), MH-55489 (to the University of Texas), and MH-55485 (to the Medical University of South Carolina).

References

1. Harris EC, Barracloough B: Suicide as an outcome for mental disorders. Br J Psychiatry 1997; 170:205–228

TABLE 3. Demographic and Baseline Characteristics of Patients With Baseline 24-Item Hamilton Depression Rating Scale Item 3 Scores ≥3 by Whether or Not Rating Resolved to 0 by End of Acute ECT Course

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients With 0 Rating (N=106)</th>
<th>Patients Without 0 Rating (N=25)</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>Mean 48.5 SD 13.8</td>
<td>Mean 41.6 SD 12.8</td>
<td>t = -2.3 df 129 p &lt;0.02</td>
</tr>
<tr>
<td>Baseline Hamilton depression scale score</td>
<td>Mean 38.6 SD 7.8</td>
<td>Mean 38.8 SD 5.3</td>
<td>t = 0.1 df 129 p 0.90</td>
</tr>
<tr>
<td>Number of ECT sessions</td>
<td>Mean 7.1 SD 3.2</td>
<td>Mean 9.1 SD 3.0</td>
<td>t = 2.8 df 129 p &lt;0.005</td>
</tr>
<tr>
<td>Psychosis</td>
<td>30 N 28.3 %</td>
<td>5 N 20.0 %</td>
<td>χ² = 0.7 df 1 0.46</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>96 N 90.6 %</td>
<td>24 N 96.0 %</td>
<td>χ² = 1.1 df 3 0.77</td>
</tr>
<tr>
<td>African American</td>
<td>6 N 5.7 %</td>
<td>1 N 4.0 %</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td>37.9 χ² = 2 df &lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remitter</td>
<td>81 N 76.4 %</td>
<td>3 N 12.0 %</td>
<td></td>
</tr>
<tr>
<td>Nonremitter</td>
<td>8 N 7.6 %</td>
<td>10 N 40.0 %</td>
<td></td>
</tr>
<tr>
<td>Dropout</td>
<td>17 N 16.0 %</td>
<td>12 N 48.0 %</td>
<td></td>
</tr>
</tbody>
</table>

a Independent sample (pooled test) comparing means.

b Comparing proportions (%).
ECT FOR SUICIDAL INTENT

Isotretinoin (13-cis-retinoic acid) is a retinoid that inhibits sebaceous gland functioning, keratinization, and inflammatory responses and is currently approved by the Food and Drug Administration (FDA) for the treatment of cystic acne (1, 2). Isotretinoin has been prescribed to 2 million patients in the United States and over 8 million patients worldwide and is highly effective for acne. The exact mechanism of action of isotretinoin remains unknown.

In the last several years there has been controversy over the possible role of isotretinoin in the development of depression and suicide (3–6). Case reports in the literature describe depression that developed in conjunction with isotretinoin treatment, resolved with discontinuation of the medication, and in some cases returned when the medication was restarted (7–14). Estimates of the incidence of depression following treatment with isotretinoin include 1% (14), 4% (15), and 6% (10). Other reports have noted suicidality, behavioral disturbances, and psychotic-type symptoms with isotretinoin treatment in addition to the typical symptoms of depression (8, 12). Isotretinoin is chemically similar to the retinoid vitamin A, a fat-soluble vitamin stored in high concentrations in the liver. Vitamin A is converted after oxidation to retinoic acid, when it has biological effects. Arctic explorers who fed on polar bear liver developed symptoms of confusion and psychosis. Large doses of vitamin A can have a number of other neurological and mental effects, including fatigue, decreased interest, headache, and diplopia (double vision) (16, 17). Published case reports of vitamin A toxicity include symptoms of aggression, personality changes, depression, poor concentration, tearfulness, psychotic symptoms, and guilty ruminations (17–19) that resolved with discontinuation of vitamin A. Among reports to the World Health Organization and the FDA of adverse events associated with acne treatments, adverse events related to depression and suicide have been more common with isotretinoin than with other treatments for acne, such as antibiotics (4, 11).

The relationship between isotretinoin treatment and depression and suicide, however, remains controversial. Although the manufacturer, on the basis of FDA guidelines, lists depression as a possible side effect, there is no consensus on a causal role for isotretinoin in the development of depression and suicide. The high incidence of depression in the general population makes it difficult to identify small increases specifically related to an additional factor, such as isotretinoin administration. One large epidemiological study did not demonstrate a significantly increased risk for suicide in patients treated with isotretinoin (3). Some authors have argued that cases of depression associated with isotretinoin administration are merely coincidental (20) or that isotretinoin actually leads to an improvement in anxiety and depression because of the clearing of disfiguring acne (21). Studies have shown an improvement in feelings of general well-being or self-image (22, 23) or in feelings of anxiety (20, 24–28) among patients with cystic acne following isotretinoin administration, although the findings were more directly related to improvement in measures of patient satisfaction, rather than clinical symptoms of depression.
To establish a causal role of isotretinoin in the development of depression and suicide, it is critical to establish a plausible biological pathway. This requires that isotretinoin must enter the central nervous system (CNS) and have an effect on the functioning of brain areas and neurochemical systems that mediate depression. Retinoids have important effects on the developing brain in animal studies (29, 30), and use of isotretinoin during pregnancy has long been known to result in CNS defects in newborns (31). Multiple positron emission tomography (PET) and single photon emission computed tomography (SPECT) studies have shown low metabolism and/or blood flow at baseline in depressed subjects in the left (32–35) and bilateral (36–43) dorsolateral prefrontal cortex and medial prefrontal cortex/anterior cingulate (34, 38, 41–48) or blunted activation with cognitive tasks in the anterior cingulate (49, 50). Other PET and SPECT studies of patients with unipolar depression showed low metabolism and/or blood flow in the caudate (36–41, 51, 52), thalamus (37), temporal cortex (37, 38, 51, 53, 54), parietal cortex (34, 40, 51), and left putamen (37). Experimental induction of depression resulted in a specific decrease in metabolism in the orbitofrontal cortex (part of the prefrontal cortex) (55, 56). The purpose of the current study was to assess the effects of isotretinoin treatment on brain functioning. We hypothesized that treatment with isotretinoin, but not antibiotic, would be associated with a decrease in orbitofrontal cortical brain metabolism as measured with $^{18}$F]fluorodeoxyglucose (FDG) PET.

**Method**

**Subjects**

The study participants included 28 healthy men and women between the ages of 18 and 50 years with treatment-resistant acne, as defined by a failed 3-month antibiotic trial, who were seeking a second trial of an antibiotic or isotretinoin. Subjects were recruited by advertisement. They were not randomly assigned to treatment with isotretinoin or placebo; they had decided with their physicians to take either a second trial of an antibiotic or a trial of isotretinoin. Because of the side effects of isotretinoin (severe skin dryness) it was decided that it would not be possible to blind the subjects or the raters to treatment condition. Subjects with serious medical or neurological illness, organic mental disorder or current psychiatric illness according to the Structured Clinical Interview for DSM-IV (57), premenstrual dysphoric disorder, current alcohol or substance abuse or dependence, retained metal that would prevent magnetic resonance imaging (MRI) scanning, a history of head trauma or loss of consciousness, a history of cerebral infectious disease, or dyslexia were excluded. Postmenopausal women were excluded.

This project was approved by the Emory University Human Investigation Committee. All subjects provided written informed consent for participation. The subjects were paid for their participation.

Each subject received a PET brain scan at baseline and again after 4 months of treatment with an antibiotic (N=15) or isotretinoin (N=13). The subjects were treated by their outpatient physicians with 1 mg/kg of isotretinoin or with an antibiotic in a standard 4-month course of treatment for acne. The antibiotics used included doxycycline (N=10), minocycline (N=2), tetracycline (N=2), and erythromycin (N=1). The subjects continued treatment until they completed the second PET scan.

Psychiatric diagnoses were established with the Structured Clinical Interview for DSM-IV, Patient Edition (SCID-P) (57). There were no current psychiatric conditions in any subject according to the SCID-P. Two of the subjects in the antibiotic group had past psychiatric conditions, including a past history of major depression in one and a past history of bulimia and alcohol dependence in another. None of the isotretinoin subjects had a current or past psychiatric disorder. None of the subjects had current alcohol or substance abuse or dependence.

**Behavioral Assessment**

Symptoms of depression were measured by using the Hamilton Depression Rating Scale at baseline and every month after the initiation of treatment (58). The severity of acne was measured with a clinician-administered acne questionnaire before and after treatment, on a scale of 0 (no acne) to 6 (very bad acne). The patients’ subjective evaluations of the severity of acne on the face and back and their feelings of depression related to their acne, on a scale of 0 (not at all) to 4 (very severe), were recorded before and after treatment by using the Skindex questionnaire (59). The Skin- dex is a 16-item self-report questionnaire with questions about emotional, functional, and symptomatic aspects of acne that has been validated for use in acne patient populations. It was also administered before and after treatment. The patients were also evaluated during the course of treatment for symptoms related to treatment.

**PET and MRI Scanning**

Two PET scans of resting brain metabolism were performed 4 months apart, before and after treatment with isotretinoin or antibiotic. The PET scans took place at 11:00 a.m. The subjects were scanned with an ECAT EXACT 921 PET camera (CTI Molecular Imaging, Knoxville, Tenn.). The ECAT EXACT has an axial field of view of 16.2 cm, the total system sensitivity is 216 kcps/µCi per ml for a 20-cm cylinder phantom in two dimensions, and the approximate axial resolution is 5.0 mm (60). Each subject was placed in a preparation room adjacent to the PET scanner room, and an intravenous line was inserted in the hand and warmed with a heating pad for measurement of arterialized venous blood samples. This method has been shown to yield metabolic values equivalent to those obtained by arterial line placement (61). The subject then received an intravenous injection of 10 mCi (370 MBq) of FDG in a single bolus. Twenty-three arterialized venous blood samples were obtained at multiple time points after injection for measurements of radioactivity in the plasma, which were used for construction of a plasma time activity curve. Three blood samples were also obtained for measurement of plasma glucose concentrations. The subject was then placed in the scanner with his or her head held in a head holder to minimize patient motion. The head was positioned with the canthomeatal line parallel to the external laser light. Following positioning within the camera gantry, postinjection transmission data were collected by using rod windowing with three orbiting $^{67}$Ga/$^{68}$Ge rod sources (60). These data were used to correct the emission data for attenuation due to overlying bone and soft tissue. The subject underwent emission scanning of the brain over the 40–60 minutes after injection with his or her eyes open in a dimly lit room. Brain and tissue measurements were used to estimate the cerebral glucose metabolic rate (in milligrams per minute per 100 milliliters) (62, 63). In one patient blood samples could not be obtained, and this patient’s data were used only for the analysis of the ratio of regional metabolism to whole brain metabolism. A 20-cm cylindrical fluid-filled phantom with a known amount of radioactivity was scanned in order to obtain calibration factors for conversion of native pixel values into units of millicuries per milliliter.
MRI scans were obtained in all subjects for coregistration with the PET scans and determination of regions of interest from the MRI scans resliced to correspond to the PET slices. MRI scans in the same subjects were obtained on a 1.5-T Philips Gyroscan Intera device (Philips Medical Systems, Andover, Mass.). Axial images were acquired with a T1-weighted gradient echo three-dimensional sequence with TR=35 msec, TE=12 msec, flip angle=35°, number of excitations=2, matrix=256x256, field of view=22 cm, and slice thickness=3 mm.

**Image Processing and Analysis**

The PET and MRI scans were transferred to a workstation for analysis. A surface-matching algorithm and the ANALYZE software package (Mayo Clinic, Rochester, Minn.) were used for coregistration of images (64). Brain surfaces from PET and MRI were matched by using this program. The MRI scan was resliced to correspond to the PET slices. Using this technique (65), we have shown a registration error of 2.86 mm. Regions of interest were drawn on the resliced MRI scans by a blinded rater using specific criteria based on anatomical landmarks with a method that we have shown to be highly reliable (66). Multiple brain regions were selected for analysis, including the temporal cortex, inferior, middle, and superior frontal gyri, superior portion of the dorsolateral prefrontal cortex, thalamus, putamen, caudate, occipital cortex, subcallosal gyrus, orbitofrontal cortex, anterior cingulate, postcentral gyrus, hippocampus, amygdala, and midbrain. These regions correspond to the regions measured in our prior studies of neural findings associated with a return of depressive symptoms induced by tryptophan depletion (55) and alpha-methylparatyrosine (56), since a primary aim of the current study was to replicate the brain findings of those prior studies. Global brain metabolism was calculated as the mean of brain tissue in all slices, including gray and white matter and the ventricular spaces.

**Data Analysis**

The brain regions were separated into those that were and were not hypothesized to change with isotretinoin. The region most consistently affected in our two prior studies of depression was the orbitofrontal cortex (55, 56). This region has also been reported to be smaller in volume in depressed patients than in comparison subjects (67, 68).

The data were analyzed to determine differences in the changes from pre- to posttreatment in regional brain metabolic rates between the isotretinoin and antibiotic treatment groups, by using repeated-measures analysis of variance with time (before and after treatment) as the repeated factor and treatment status (isotretinoin versus antibiotic) and hemisphere (left versus right) as factors in the analysis. The interaction between treatment status and time was examined in this model. Secondary analysis examined the ratio of regional to whole brain metabolism, and it added baseline global metabolism as a factor in the model. Bonferroni corrections were performed to correct for multiple comparisons (p=0.05/12=0.004).

Correlations between brain metabolism and behavioral variables were also examined by comparing the relationships between the changes, from before to after treatment, in scores on the Hamilton depression scale, clinician-administered acne questionnaire, Skindex, and acne severity self-report and the change in regional brain metabolism (mean of left and right orbitofrontal metabolism) in the isotretinoin and antibiotic treatment groups. Corrections were made for multiple comparisons (p=0.05/4=0.0125). Analyses also examined the correlation between baseline orbitofrontal brain metabolism and age, education, and behavioral factors related to self-assessment of acne severity, depression, and emotions related to acne, measured with the Skindex. Corrections were made for multiple comparisons (p=0.05/6=0.008).

The data were analyzed by using the SAS System for Windows V8 (SAS, Cary, N.C.).

**Results**

**Recruitment and Demographic Factors of Subjects**

Eighty-eight subjects were initially screened for participation in the study. Of these, 44 met the criteria for participation according to the initial screening, signed informed consent statements, and were identified as acne patients.
who were beginning a second trial of antibiotic (N=22) or who were going to be treated with isotretinoin (N=22) (Figure 1). Twenty-eight subjects completed participation in this protocol, including pre- and posttreatment imaging. Of these, 13 were treated with isotretinoin and 15 with antibiotics.

Demographic, behavioral, and acne-related variables related to the two treatment groups before the initiation of treatment are presented in Table 1. The isotretinoin subjects had fewer years of education and were younger, but the latter difference was not statistically significant. They did not differ significantly from the antibiotic group in their reasons for receiving treatment (cystic acne, psychological distress, scarring, or a combination). According to the clinician ratings, the isotretinoin subjects had more severe pretreatment acne than the subjects receiving antibiotics. However, according to the self-ratings there were no differences in acne on the face or back or in feelings of depression related to the acne. There were no differences between the two groups in behavioral, emotional, and functional effects of acne as measured by the Skindex. There were also no differences in baseline depressive symptom levels as measured by the Hamilton depression scale.

### Effects of Isotretinoin on Regional Cerebral Metabolism

Administration of isotretinoin but not antibiotic was associated with decreased brain metabolism in the orbitofrontal cortex after 4 months of treatment (Figure 2, Figure 3). This effect was seen for both absolute metabolism (Figure 2, Table 2) and for the ratio of orbitofrontal to whole brain metabolism (F=4.64, df=1, 110, p<0.05). A secondary analysis included pretreatment whole brain metabolism in the model and also showed greater reductions in orbitofrontal metabolism after treatment in the isotretinoin group than in the antibiotic group (F=9.66, df=1, 104, p=0.002). The value for the interaction between treatment status (isotretinoin versus antibiotic) and time (before and after treatment) and the percentage change in the mean metabolic value with treatment are presented for each region in Table 2. Differences in functioning between the groups at the p<0.05 level were also seen for the middle frontal gyrus, thalamus, hippocampus, caudate, and putamen. These differences were not significant after correction for multiple comparisons, however, and there were no differences after we corrected for whole brain metabolism by examining differences in the ratios of regional to whole brain values.

The mean pretreatment rate of metabolism in the orbitofrontal cortex was higher for patients in the isotretinoin group than for those in the antibiotic group (F=2.05, df=7, 107, p=0.03). This was not hypothesized a priori and was not significant after correction for multiple comparisons.

### Relationship Between Behavior and Brain Metabolism

Five patients treated with isotretinoin had symptoms of headache. These patients also had subtle changes in irritability and/or mood as assessed by self, family, or the research staff. These subjects all had decreases in brain metabolism with isotretinoin administration (Figure 2). A representative subject is shown in Figure 3. However, these subjects did not show clinically significant depres-
Am J Psychiatry 162:5, May 2005

**BREMNER, FANI, ASHRAF, ET AL.**

http://ajp.psychiatryonline.org

sion as assessed with the Hamilton depression scale (Figure 4). One subject in the isotretinoin group and one in the antibiotic group had a clinically significant increase in depression as measured by the Hamilton scale (greater than 9-point increase); however, there were no significant increases in Hamilton depression scores in the groups as a whole and no significant differences between groups.

Patients in the isotretinoin group had more severe acne as rated by clinician assessment at baseline (F=18.80, df=1, 25, p<0.05), and they had a greater improvement with isotretinoin treatment, as indicated by less severe acne according to the clinician-administered questionnaire after treatment than before treatment, than did patients receiving antibiotics (F=22.73, df=1, 25, p<0.05). There was also a greater improvement with isotretinoin in self-reported acne (F=2.62, df=4, 88, p<0.05). There were no differences in change in feelings of “depression related to acne” between the groups. There was no relationship in either the isotretinoin or antibiotic group between baseline orbitofrontal cortical metabolism and depression as measured with the Hamilton scale, self-reported or clinician-assessed acne severity as measured with the analogue scales, feelings of depression related to acne as measured with the analogue scale or the Skindex, or overall psychological effects of acne as measured with the Skindex. The decrease in orbitofrontal cortical metabolism with treatment in the isotretinoin group was correlated with a single item of the Skindex at baseline, effect of acne on ability to work (r=-0.67, N=9, p=0.03). This correlation was not significant after correction for multiple comparisons. There was no relationship between “worrying about skin condition” as measured with the Skindex and orbitofrontal metabolism at baseline or with treatment. There was no correlation between acne-related depression at baseline and decrease in orbitofrontal cortical metabolism with isotretinoin.

There was no correlation between age or education and baseline orbitofrontal cortical metabolism or change in metabolism with treatment.

**Discussion**

A 4-month treatment trial with isotretinoin was associated with a decrease in brain functioning in the orbitofrontal cortex, a brain region implicated in depression. These changes were not seen after a similar course of treatment with an antibiotic. After correction for differences in whole brain metabolism, this effect was specific to the orbitofrontal cortex. The greatest magnitude of decrease was observed in subjects who developed symptoms of headache during the course of treatment with isotretinoin. Isotretinoin was not associated, however, with any changes in depressive symptom severity as measured with the Hamilton depression scale.

Isotretinoin has a variety of effects on brain neurochemical systems (69–71). Retinoids modulate gene expression in the brain in a broad spectrum and have effects on several neurochemical systems, including the dopamine sys-

---

**FIGURE 2. Effects of Isotretinoin and Antibiotics on Orbitofrontal Cortical Metabolism in Patients Receiving Treatment for Acne**

<table>
<thead>
<tr>
<th>Glucose Metabolic Rate (mg/min per 100 g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Treated With Isotretinoin</td>
</tr>
<tr>
<td>Subject Treated With Antibiotics</td>
</tr>
</tbody>
</table>

![](image1.png)

---

**FIGURE 3. Effects on Regional Brain Metabolism in a Representative Patient Receiving Isotretinoin Treatment for Acne**

![](image2.png)

---

* Isotretinoin but not antibiotic administration resulted in a significant decrease in orbitofrontal cortical metabolism (p<0.001, ANOVA). The mean percentage change with treatment within individual subjects was −21% for isotretinoin and 2% for antibiotic.
Isotretinoin and Brain Functioning

High levels of the enzyme involved in retinoid synthesis, aldehyde dehydrogenase, are found in mesostriatal and mesolimbic dopamine pathways (72, 73). Dopamine mesocortical pathways involve release of dopamine transmitter in the orbitofrontal cortex and other parts of the prefrontal cortex. Isotretinoin may influence these pathways. Administration of retinoids causes changes in dopamine receptors (74), while genetic mutations of retinoid receptors are associated with deficits in dopamine receptors as well as mesolimbic dopamine functioning (75). Retinoids are associated with an inhibition of neurogenesis in the hippocampus (76), a brain area with connections to prefrontal cortical areas, including the orbitofrontal cortex. Inhibition of neurogenesis in the hippocampus has been hypothesized to play a role in depression (77–80). Retinoids also have effects on brain trophic factors (81). These findings have led to the hypothesis that retinoids play a role in the development of psychiatric disorders (69, 82).

A number of limitations of the current study are worthy of mention. This was a pilot study designed to evaluate the possibility of an effect of isotretinoin on brain functioning. For this reason the study group was small, which may have contributed to the fact that we did not observe treatment-related changes in mood as assessed by the behavioral ratings in this study. Some patients, however, complained of headache with isotretinoin, and these patients exhibited greater decreases in orbitofrontal brain metabolism during isotretinoin treatment. Because of the costs of isotretinoin we were unable to pay for this medication for all subjects. Therefore, we were unable to randomly assign subjects to treatment with isotretinoin or antibiotic and were unable to control which antibiotic the subjects were taking. We therefore recruited subjects who were preparing to undergo a second treatment course with an antibiotic or to switch to isotretinoin on the basis of a decision made in conjunction with the subject’s own physician. This method likely contributed to the fact that the isotretinoin subjects had more severe acne at baseline. The isotretinoin group also had less education. We examined a variety of demographic factors, including age, education, psychological distress, and self- and clinician assessments of acne severity, and found no relationships with baseline orbitofrontal metabolism or change in orbitofrontal functioning with treatment. An unexpected finding was a pat-

### TABLE 2. Regional Brain Metabolic Rates of Patients Before and After Isotretinoin or Antibiotic Treatment for Acne

<table>
<thead>
<tr>
<th>Brain Region</th>
<th>Isotretinoin (N=13)</th>
<th>Antibiotic (N=15)</th>
<th>Interaction of Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rate (mg/minute per 100 g)</td>
<td>Rate (mg/minute per 100 g)</td>
<td>Group and Time</td>
</tr>
<tr>
<td></td>
<td>Baseline Mean SD</td>
<td>After Treatment Mean SD</td>
<td>% Changea</td>
</tr>
<tr>
<td>Orbitofrontal cortex</td>
<td>12.73 3.59 10.70 2.86</td>
<td>–16</td>
<td>10.07 3.08 10.87 3.30</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>14.14 4.82 14.11 3.84</td>
<td>0</td>
<td>12.29 4.25 14.39 4.04</td>
</tr>
<tr>
<td>Superior dorsolateral prefrontal cortex</td>
<td>13.88 4.55 13.83 3.80</td>
<td>0</td>
<td>12.64 4.26 13.67 3.93</td>
</tr>
<tr>
<td>Thalamus</td>
<td>10.39 2.71 10.00 2.52</td>
<td>–4</td>
<td>8.85 2.36 10.19 2.07</td>
</tr>
<tr>
<td>Temporal cortex</td>
<td>12.90 4.64 12.89 3.73</td>
<td>0</td>
<td>11.85 4.55 12.96 4.13</td>
</tr>
<tr>
<td>Superior frontal cortex</td>
<td>13.01 4.02 13.49 5.75</td>
<td>4</td>
<td>11.86 3.59 13.06 3.32</td>
</tr>
<tr>
<td>Inferior frontal cortex</td>
<td>14.05 4.78 13.23 3.98</td>
<td>–6</td>
<td>12.33 4.83 13.56 4.23</td>
</tr>
<tr>
<td>Caudate</td>
<td>12.52 3.81 12.43 3.28</td>
<td>–1</td>
<td>11.08 2.73 12.61 3.27</td>
</tr>
<tr>
<td>Putamen</td>
<td>13.67 3.52 13.86 3.83</td>
<td>1</td>
<td>12.34 3.52 14.07 3.41</td>
</tr>
<tr>
<td>Anterior cingulate</td>
<td>11.58 4.00 11.21 2.83</td>
<td>–3</td>
<td>10.31 3.23 11.26 3.83</td>
</tr>
<tr>
<td>Postcentral gyrus</td>
<td>13.63 4.82 13.27 3.53</td>
<td>–3</td>
<td>12.22 4.03 13.34 3.90</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>8.09 3.12 7.80 2.29</td>
<td>–4</td>
<td>7.55 2.69 8.65 2.26</td>
</tr>
<tr>
<td>Subcallosal region</td>
<td>10.39 2.40 11.72 3.65</td>
<td>13</td>
<td>10.83 2.58 11.09 3.05</td>
</tr>
<tr>
<td>Amygdala</td>
<td>8.20 3.10 7.83 2.97</td>
<td>5</td>
<td>7.68 2.34 8.32 1.78</td>
</tr>
<tr>
<td>Midbrain</td>
<td>7.12 2.04 6.49 1.69</td>
<td>–9</td>
<td>6.29 1.86 6.94 1.51</td>
</tr>
<tr>
<td>Global</td>
<td>11.13 3.47 10.53 2.66</td>
<td>–5</td>
<td>9.86 2.95 11.02 3.01</td>
</tr>
</tbody>
</table>

a Percent change in mean metabolic rate.

#### FIGURE 4. Effects of Isotretinoin and Antibiotic on Symptoms of Depression in Patients Receiving Treatment for Acne

Depression was measured with the 21-item Hamilton Depression Rating Scale. There was no significant effect of either isotretinoin or antibiotic on depressive symptom severity.

---

a Depression was measured with the 21-item Hamilton Depression Rating Scale. There was no significant effect of either isotretinoin or antibiotic on depressive symptom severity.
tern of greater baseline functioning in the orbitofrontal cortex in the isotretinoin group than in the antibiotic group. One might question whether factors in the isotretinoin group, such as increased worrying related to more severe acne, might have contributed to differences in brain functioning. Obsessive-compulsive disorder has been associated with higher orbitofrontal metabolism (83). However, we did not observe any differences in depression, psychological distress, or even self-rating of acne severity between the two groups before treatment. We also did not find a correlation between orbitofrontal metabolism at baseline or change with treatment and the item related to “worrying about skin condition” on the Skindex. Antibiotics act by inhibiting bacterial protein synthesis. Although they can pass the blood-brain barrier, they are not known to have effects on brain functioning. The most common side effects of antibiotics are nausea, vomiting, and diarrhea (84), and psychiatric and neurological side effects are much more common in acne patients treated with isotretinoin than those treated with antibiotics. For this reason it is unlikely that effects of antibiotics on brain functioning would account for the results of the current study. We excluded subjects with a history of mental illness. This may have involved exclusion of subjects who were prone to the development of depression and may limit the generalizability of the findings. In summary, a randomized, placebo-controlled study would provide more definitive results than the current study.

To our knowledge, this is the first study of the effects of isotretinoin on human brain functioning. The findings suggest that isotretinoin may affect brain functioning, providing a possible biological mechanism by which isotretinoin treatment could lead to depression in a minority of vulnerable acne patients. Future studies using randomized designs to evaluate the effects of isotretinoin on brain functioning are warranted.

References

27. Rubinow DR, Peck GL, Quallace KM, Gantt GG: Reduced anxiety and depression in cystic acne patients after successful treat-
ISOTRETINOIN AND BRAIN FUNCTIONING


69. Goodman AB: Three independent lines of evidence suggest retinoids as causal to schizophrenia. Proc Natl Acad Sci USA 1998; 95:7240–7244

70. Jentsch JD, Roth RH, Taylor JR: Role for dopamine in the behavioral functions of the prefrontal corticostriatal system: implications for mental disorders and psychotropic drug action. Prog Brain Res 2000; 126:433–453


82. Goodman AB: Chromosomal locations and modes of action of genes in the retinoid (vitamin A) system support their involvement in the etiology of schizophrenia. Am J Med Genet Neuropsychiatry Genet 1995; 60:335–348


Higher Levels of Basal Serial CSF Cortisol in Combat Veterans With Posttraumatic Stress Disorder

Dewleen G. Baker, M.D.
Nosa N. Ekhator, M.S.
John W. Kasckow, M.D.
Boris Dashevsky, Ph.D.
Paul S. Horn, Ph.D.
Ludmila Bednarik, M.S.
Thomas D. Geracioti, Jr., M.D.

Objective: Results of basal peripheral cortisol measures in posttraumatic stress disorder (PTSD) have been variable. The authors’ goal was to measure CSF cortisol concentrations, which more accurately reflect brain glucocorticoid exposure, in subjects with or without PTSD.

Method: CSF was withdrawn from a subarachnoid catheter and plasma from a venous catheter, both indwelling, over a 6-hour interval to determine hourly plasma ACTH and cortisol concentrations and hourly CSF cortisol levels in eight well-characterized combat veterans with PTSD and eight matched healthy volunteers.

Results: Mean CSF cortisol concentrations were significantly higher in the subjects with PTSD (3.18 ng/ml, SD=0.33) than in the normal volunteers (2.33 ng/ml, SD=0.50), largely due to higher CSF cortisol concentration nadirs. No group differences were observed in either plasma ACTH or peripheral (plasma or urinary free) cortisol. CSF corticotropin-releasing hormone and CSF cortisol concentrations were positively and significantly correlated.

Conclusions: Despite normal peripheral cortisol indexes in the veterans with PTSD, their CNS exposure to cortisol was greater than that of normal comparison subjects.

Circadian studies have found low peripheral basal cortisol concentrations in posttraumatic stress disorder (PTSD), whereas results of single time-point plasma and urinary free cortisol studies have been variable (1, 2). Given the elevation in CSF corticotropin-releasing hormone (CRH) in PTSD, even normal circulating cortisol levels could be considered low (3). Cortisol enters the brain readily from plasma, but levels in the CSF cannot be inferred directly from peripheral measures (4–7). CSF cortisol concentrations, however, have never been examined in PTSD. We used serial CSF sampling to extend the study of cortisol in PTSD into the CNS.

Method

Patients and Normal Volunteers

The study was approved by the Institutional Review Board of the University of Cincinnati Medical Center and the Research Committee of the Cincinnati Veterans Affairs Medical Center. Written informed consent was obtained from all participating subjects.

We studied 16 medication-free subjects—eight healthy men and eight men with combat-related PTSD, interviewed with the Structured Clinical Interview for DSM-III-R. We excluded current or past psychiatric disorder or current substance abuse (except tobacco) in healthy volunteers and a history of these disorders in their first-degree relatives. The mean ages of the healthy volunteers and PTSD subjects were 41.3 years (SD=8.7) and 41.0 years (SD=10.4), respectively, and the mean body mass indexes of the volunteers and PTSD patients were 25.4 kg/m² (SD=4.0) and 29.3 kg/m² (SD=4.2), respectively. All PTSD subjects had been exposed to severe combat trauma, and all were free of current substance abuse (except tobacco) (data on file). The patients’ mean Clinician Administered PTSD Scale and Hamilton Depression Rating Scale scores were 81.9 (SD=23.2) and 12.6 (SD=4.3), respectively.

Subjects were admitted to the Clinical Research Unit the day before the CSF withdrawal procedure, according to a standard protocol (3, 8). A 20-gauge catheter was placed in the lumbar subarachnoid space at 8:00 a.m. the morning of the withdrawal, and continuous CSF withdrawal into iced test tubes was begun approximately 3 hours after catheter placement (3, 8). Blood was withdrawn at intervals from a heparin lock.

CSF cortisol, plasma cortisol, and ACTH were assayed by radioimmunoassay (ICN Pharmaceuticals, Costa Mesa, Calif.). Hourly CSF cortisol samples were assayed, in duplicate, as described elsewhere (9). Respective intraassay and interassay coefficients of variation were 7.2% and 2.5% for cortisol and 11.0% and 6.0% for ACTH. Previous assays of urinary free cortisol and CSF CRH were reported (3).

Repeated-measures analyses were conducted according to a linear mixed model where subjects, nested within subgroups, were treated as random effects while all other covariates (time, age, body mass index) were treated as fixed effects. Pearson correlation coefficients were used in all correlation analyses, and t tests were performed on means; p values less than 0.05 were considered significant.

Results

CSF cortisol concentrations were significantly higher in the patients with PTSD (mean=3.18 ng/ml, SD=0.33) than in the normal volunteers (mean=2.33 ng/ml, SD=0.50) (Figure 1). Using repeated-measures analysis, we found a main effect for group (F=9.81, df=1, 12, p<0.01) but no significant group-by-time interaction. The coefficient of variation (standard deviation/mean) calculated for each study subject across the six time points was significantly lower.
for the patients with PTSD than for the healthy volunteers (F=15.37, df=1, 11, p<0.003). Additionally, a comparison of zenith versus nadir CSF cortisol concentrations over the collection period revealed a significant group (PTSD versus healthy) difference in CSF minimum cortisol levels (t=5.74, df=14, p<0.0001) but not in maximum cortisol levels (t=1.06, df=14, n.s.) (Figure 1).

Mean plasma ACTH concentrations for the patients with PTSD and the normal volunteers were 108.49 pg/ml (SD=63.01) and 81.90 ng/ml (SD=17.96), respectively; mean plasma cortisol concentrations for the patients and volunteers were 78.26 pg/ml (SD=66.44) and 67.02 ng/ml (SD=27.12), respectively. No group, time, or interaction effects for either plasma ACTH (F=1.41, df=1, n.s.) or cortisol (F=0.94, df=1, n.s.) were found, nor were the group mean 24-hour urinary free cortisol excretions significantly different (χ²=0.013, df=1, n.s.).

There was a trend for a correlation between CSF and plasma cortisol concentrations for all experimental subjects taken together (r=0.50, N=15, p<0.10). (Information on cortisol concentrations was missing for one volunteer in the correlation analysis.) The mean CSF-to-plasma cortisol ratios were 4.45% (SD=1.47%) (all subjects), 4.46% (SD=1.19%) (patients with PTSD), and 4.44% (SD=1.84%) (healthy volunteers), with no significant group differences. However, the CSF-to-urinary free cortisol ratio tended to be higher in the PTSD group (mean=3.94%, SD=2.38%, for all participants; mean=4.61%, SD=3.02%, for patients with PTSD; and mean=3.36%, SD=1.65%, for volunteers) (F=3.99, df=1, 11, p<0.08).

Neither plasma cortisol nor ACTH mean concentrations showed any significant relation to CSF CRH levels. However, mean CSF cortisol concentrations were significantly correlated with CSF CRH for all subjects (r=0.624, N=16, p<0.05) and showed similar, but nonsignificant, correlations within the PTSD group (r=0.47, N=8, n.s.) and the volunteer group (r=0.54, N=8, n.s.) (Figure 1).

**Discussion**

Higher basal CSF cortisol concentrations were observed in veterans with chronic PTSD than in healthy subjects, despite normal plasma cortisol and 24-hour urinary free cortisol concentrations. The higher CSF cortisol concentrations were uniformly attributable to higher CSF cortisol nadirs in the PTSD subjects over the 6-hour collection period; CSF cortisol zeniths were similar across study groups. Accordingly, intradividual variability in the CSF cortisol concentration was diminished in the men with chronic PTSD relative to healthy volunteers.

The significant positive correlation observed between mean CSF CRH and CSF cortisol levels may reflect a bidirectional regulation of glucocorticoids and CRH. Although the ratio of CSF to plasma cortisol was nearly identical in both groups, the ratio of CSF cortisol to urinary free cortisol was higher in the PTSD group than the healthy group. Cortisol, unbound in urine, is mostly free (<10% bound) in CSF and largely protein-bound (>90%) in plasma (4, 10). Studies show that the blood-CSF steroid interchange is complex; CSF glucocorticoid levels respond to increases in peripheral cortisol, but other, as yet poorly understood, factors independently modulate CNS cortisol levels (5, 7, 11, 12). Multidrug-resistant P-glycoprotein regulates CNS cortisol access (6). Preclinical studies demonstrating the CNS presence of precursors and enzymes for cortisol synthesis raise the possibility of de novo brain synthesis in humans (13). Therefore, the elevated ratio of CSF to urinary free cortisol excretion in the PTSD group could be attributable to 1) an increase in unbound cortisol available for transport or differences in rate of transport into the CNS, 2) a decrease in CSF cortisol clearance or metabo-
lism, 3) increased de novo cortisol synthesis in the CNS, or 4) diminished tissue uptake of CSF cortisol by the brain.

References


Childhood Trauma and Personality Disorder: Positive Correlation With Adult CSF Corticotropin-Releasing Factor Concentrations

Royce Lee, M.D.
Thomas D. Geracioti, Jr., M.D.
John W. Kasckow, M.D., Ph.D.
Emil F. Coccaro, M.D.

Objective: To test the hypothesis that early life trauma results in adult stress hormone alterations in individuals with personality disorders, the authors examined the relationship between history of childhood adversity and lumbar CSF corticotropin-releasing factor (CRF).

Method: Participants were 20 otherwise healthy men who met DSM-IV criteria for personality disorders. CSF CRF was obtained by lumbar puncture, and childhood adversity was measured by the Childhood Trauma Questionnaire. Correlations were obtained between CSF CRF and the total score on the Childhood Trauma Questionnaire as well as scores on its four subscales.

Results: CSF CRF concentrations were positively correlated with the total score on the Childhood Trauma Questionnaire. Analysis of the subscales revealed that CSF CRF was correlated with emotional neglect. Correlations between CSF CRF level and physical and emotional abuse and with physical neglect were not statistically significant.

Conclusions: Consistent with the hypothesis that the severity of early life stress is correlated with stress hormone abnormalities in adulthood, Childhood Trauma Questionnaire total scores and emotional neglect scores were significantly correlated with CSF CRF levels in individuals with personality disorders.

(Tried at 2005; 162:995–997)
Results

CSF CRF levels were significantly, positively correlated with Childhood Trauma Questionnaire total scores (r=0.48, N=20, p<0.04) (Figure 1). Median split analysis revealed the same relationship: CSF CRF levels of subjects with scores above the median Childhood Trauma Questionnaire total score were significantly higher than those of subjects with scores below the median total score (mean=60.5, SD=36.3, versus mean=28.5, SD=14.2) (t=2.60, df=18, p<0.03). Subsequent analysis of the four Childhood Trauma Questionnaire subscales revealed a statistically significant positive correlation between CSF CRF levels and emotional neglect (r=0.45, N=20, p<0.05). Correlations between CSF CRF concentration with physical and emotional abuse and physical neglect were smaller in magnitude and not significant (r=0.33, N=20, p=0.15, and r=0.34, N=20, p<0.14, respectively). The relationship between CSF CRF and sexual abuse could not be examined because of the restriction of range in scores for this Childhood Trauma Questionnaire subscale. Age, race, socioeconomic status, and GAF were not correlated with CSF CRF levels or with Childhood Trauma Questionnaire total score.

Despite the fact that the CSF CRF level correlated significantly with CSF ACTH level (r=0.46, N=20, p<0.05), no correlation was noted between CSF ACTH and the total Childhood Trauma Questionnaire score. CSF CRF levels were not correlated with any of the personality variables measured (Tridimensional Personality Questionnaire and Eysenck Personality Questionnaire) and were not correlated with subjective mood states as measured by visual analogue scale.

Discussion

The results of this study reveal a direct correlation between CSF CRF levels and history of childhood trauma in men with personality disorders. The results suggest that all forms of childhood trauma assessed, childhood emotional neglect has the strongest relationship with CRF levels in personality disorders. Most human studies so far have focused on other forms of trauma, such as sexual and physical abuse. However, the type of child trauma measured by the Childhood Trauma Questionnaire emotional neglect subscale is analogous to the disruptions in maternal care administered in rodent handling and separation paradigms (7) and primate variable foraging paradigms (8). These studies have found evidence that early life stress leads to greater hypothalamic-pituitary-adrenal (HPA) axis function by means of less negative feedback at the level of the hippocampus and greater CRF concentration (7, 9).

CSF CRF levels have recently been found to be positively associated with perceived early life stress in depressed and nondepressed subjects (10). Human studies have demonstrated abnormal HPA axis function in association with

FIGURE 1. History of Childhood Trauma and CSF Corticotropin-Releasing Factor Concentrations in 20 Men With Personality Disordersa

a On the left, CSF corticotropin-releasing factor (CRF) concentrations were directly correlated with the Childhood Trauma Questionnaire emotional neglect subscale scores (r=0.49, N=20, p<0.03). The box plot on the right diagrams CSF CRF concentrations in subjects divided by median split of Childhood Trauma Questionnaire total scores. Mean CSF CRF levels were significantly higher in subjects with higher total scores (mean=60.5, SD=36.3) than in those with lower total scores (mean=28.5, SD=14.2) (t=2.60, df=18, p<0.03).
history of childhood abuse in the form of enhanced ACTH hormone response to exogenous CRF (11).

Given the presence of CRF receptors in areas of the brain associated with emotional and cognitive processing such as the amygdala and neocortex, CRF may play a role in personality psychopathology. Preliminary studies have found relationships between HPA axis abnormalities and dimensional aspects of personality psychopathology such as elevated arginine vasopressin levels in impulsive-aggressive subjects (12).

The limitations of our study include a small number of subjects, inadequate power to explore the relationship between CRF and dimensional measures of temperament, the inclusion of only male subjects, retrospective assessment of history of childhood trauma, and the stressful nature of the lumbar puncture. Further work in this area should clarify the links between childhood trauma, CNS neurochemistry, and personality psychopathology.

Received June 5, 2003; revision received March 16, 2004; accepted May 26, 2004. From the Clinical Neuroscience and Psychopharmacology Research Unit, Department of Psychiatry, University of Chicago; Cincinnati VA Medical Center, Psychiatry Service; Department of Psychiatry, University of Cincinnati College of Medicine; University of Cincinnati Neurosciences Program; and Department of Psychiatry, Columbia University, New York. Address correspondence and reprint request to Dr. Lee, Clinical Neuroscience and Psychopharmacology Research Unit, Department of Psychiatry, The University of Chicago, 5841 S. Maryland Ave., Chicago, IL 60637; rlee@yoda.bsd.uchicago.edu (e-mail).

References
Circadian Rhythm of Salivary Cortisol in Holocaust Survivors With and Without PTSD

Rachel Yehuda, Ph.D.
Julia A. Golier, M.D.
Shira Kaufman, B.A.

Objective: The authors’ goal was to determine whether cortisol circadian rhythm alterations observed in younger subjects with posttraumatic stress disorder (PTSD) are also present in geriatric trauma survivors with PTSD.

Method: Salivary cortisol levels were measured at six intervals from awakening until bedtime in 23 Holocaust survivors with PTSD, 19 Holocaust survivors without PTSD, and 25 subjects who had not been exposed to the Holocaust. Thirty-three of the subjects were men, and 34 were women.

Results: Cortisol levels were significantly lower at awakening, at 8:00 a.m., and at 8:00 p.m. in Holocaust survivors with PTSD than in nonexposed subjects, resulting in a flatter circadian rhythm, similar to what has been observed in aging but different from what has been reported in younger subjects with PTSD.

Conclusions: These data provide evidence of differential neuroendocrine alterations in geriatric PTSD.

(Am J Psychiatry 2005; 162:998–1000)
sis of covariance of the six samples demonstrated a main effect of group (Pillai’s F=2.91, df=12, 114, p=0.002); differences were apparent at awakening (F=3.32, df=2, 61, p=0.04), 8:00 a.m. (F=3.89, df=2, 61, p=0.03), and 8:00 p.m. (F=3.34, df=2, 61, p=0.04) (raw data in legend of Figure 1). The differences, according to post hoc pairwise comparisons using Bonferroni adjustment, were attributable to lower cortisol at awakening and 8:00 a.m. and higher cortisol levels at 8:00 p.m. in Holocaust survivors with PTSD than in nonexposed subjects. There were no significant differences between the Holocaust survivors without PTSD and the nonexposed subjects. Because the cortisol values were skewed and kurtotic, the analyses were repeated with log-transformed cortisol levels. A similar pattern emerged: there was a main effect of group (F=3.75, df=12, 114, p=0.0055), with significant post hoc differences at awakening and 8:00 a.m. but not at 8:00 p.m. There was no significant group difference in area under the curve for the raw or log-transformed cortisol data.

Age was not significantly correlated with cortisol at any time point for the whole study group. There was a significant association between age and the raw (r=0.38, N=34, p=0.027) and log-transformed (r=0.34, N=34, p=0.046) 8:00 a.m. cortisol level in women.

Discussion

The data show that Holocaust survivors with PTSD show a flatter circadian rhythm than Jewish adults not exposed to the Holocaust. To our knowledge, this finding provides the first evidence of subtle HPA axis alterations in older people with PTSD. To the extent that these alterations are similar to those described in normal aging, the findings add further evidence to the possibility that there may be an acceleration of the aging process in PTSD. The subjects evaluated in the current study were in the early stages of aging, which may explain why the observed alterations in PTSD were modest and why an overall effect of age in this study group as a whole could not be demonstrated.

It should be noted that the sampling of saliva during waking hours may provide different information than evaluating cortisol levels across the entire diurnal cycle. In a study of 24-hour urine collections, survivors with PTSD were found to have lower cortisol levels (9), but no differences in the area under the curve for cortisol excretion were observed in our study. A previous circadian study in veterans with PTSD conducted over 24 hours, which included time periods that were not fully sampled in our study, found lower cortisol levels that appeared to result from lower nocturnal cortisol release (4).

Group differences in basal cortisol levels have not been as consistently observed in studies of women with PTSD as they have in men (see, e.g., reference 2 for review). Whereas the evaluation of gender differences in PTSD is often confounded by the fact that men and women are exposed to different types of events and at different ages, in Holocaust survivors the same types of traumatic events affected both men and women, allowing for a more direct comparison of the effects of gender. In the present study, women had higher cortisol levels at 8:00 a.m. than men, regardless of PTSD diagnosis, but there were no gender differences related to PTSD. These findings, together with our previous observations, suggest that there are no gender differences related to PTSD in basal cortisol levels in our study group. A relationship between age and cortisol was present only in women, consistent with other observations of gender differences in early aging.

In sum, this study provides preliminary evidence for subtle neuroendocrine alterations in older people with PTSD. As such, the present study offers a rationale for performing a more comprehensive circadian rhythm analysis in older subjects with PTSD and for further assessing the impact of aging on the biology of PTSD.
Received July 7, 2003; revision received Feb. 24, 2004; accepted May 24, 2004. From Mount Sinai School of Medicine and Bronx VA Medical Center. Address correspondence and reprint requests to Dr. Yehuda, Psychiatry OOMH, Bronx VA Medical Center, 130 W. Kingsbridge Rd., Bronx, NY 10468; rachel.yehuda@med.va.gov (e-mail).

References

Brief Report

Posttraumatic Stress Disorder and Memory Problems After Female Genital Mutilation

Alice Behrendt, Dipl.-Psych.
Steffen Moritz, Ph.D.

Objective: This pilot study investigated the mental health status of women after genital mutilation. Although experts have assumed that circumcised women are more prone to developing psychiatric illnesses than the general population, there has been little research to confirm this claim. It was predicted that female genital mutilation is associated with a high rate of posttraumatic stress disorder (PTSD).

Method: The psychological impact of female genital mutilation was assessed in 23 circumcised Senegalese women in Dakar. Twenty-four uncircumcised Senegalese women served as comparison subjects. A neuropsychiatric interview and further questionnaires were used to assess traumatization and psychiatric illnesses.

Results: The circumcised women showed a significantly higher prevalence of PTSD (30.4%) and other psychiatric syndromes (47.9%) than the uncircumcised women. PTSD was accompanied by memory problems.

Conclusions: Within the circumcised group, a mental health problem exists that may furnish the first evidence of the severe psychological consequences of female genital mutilation.

According to the World Health Organization definition (1), female genital mutilation (often referred to as “female circumcision”) “comprises all procedures involving the partial or total removal of the external female genitalia or other injury to the female genital organs whether for cultural, religious or other nontherapeutic reasons” (p. 25). The procedure is performed within a wide range of African ethnic groups. Appropriate estimations are that between 100 and 140 million women and girls have undergone the practice (1). Today, female genital mutilation is defined as a violation of human rights, and in many countries (e.g., Senegal, Burkina Faso) laws have been passed to outlaw female genital mutilation. Nevertheless, female genital mutilation is still performed by many African tribes. In Senegal, about 20% of the female population is estimated to be circumcised (2).

The age of the performance (of female genital mutilation) varies across and within countries. Most of the operations occur before the end of childhood (mainly between 4 and 10 years). The medical consequences have been broadly investigated. However, there has been hardly any research to qualify and quantify the impact on psychological health. Case studies mention phobias, depression, and sexual disorders (3). Nonetheless, it has been ignored
Posttraumatic Stress Disorder and Memory Problems After Female Genital Mutilation

Alice Behrendt, Dipl.-Psych.
Steffen Moritz, Ph.D.

Objective: This pilot study investigated the mental health status of women after genital mutilation. Although experts have assumed that circumcised women are more prone to developing psychiatric illnesses than the general population, there has been little research to confirm this claim. It was predicted that female genital mutilation is associated with a high rate of posttraumatic stress disorder (PTSD).

Method: The psychological impact of female genital mutilation was assessed in 23 circumcised Senegalese women in Dakar. Twenty-four uncircumcised Senegalese women served as comparison subjects. A neuropsychiatric interview and further questionnaires were used to assess traumatization and psychiatric illnesses.

Results: The circumcised women showed a significantly higher prevalence of PTSD (30.4%) and other psychiatric syndromes (47.9%) than the uncircumcised women. PTSD was accompanied by memory problems.

Conclusions: Within the circumcised group, a mental health problem exists that may furnish the first evidence of the severe psychological consequences of female genital mutilation.

According to the World Health Organization definition (1), female genital mutilation (often referred to as “female circumcision”) “comprises all procedures involving the partial or total removal of the external female genitalia or other injury to the female genital organs whether for cultural, religious or other nontherapeutic reasons” (p. 25). The procedure is performed within a wide range of African ethnic groups. Appropriate estimations are that between 100 and 140 million women and girls have undergone the practice (1). Today, female genital mutilation is defined as a violation of human rights, and in many countries (e.g., Senegal, Burkina Faso) laws have been passed to outlaw female genital mutilation. Nevertheless, female genital mutilation is still performed by many African tribes. In Senegal, about 20% of the female population is estimated to be circumcised (2).

The age of the performance (of female genital mutilation) varies across and within countries. Most of the operations occur before the end of childhood (mainly between 4 and 10 years). The medical consequences have been broadly investigated. However, there has been hardly any research to qualify and quantify the impact on psychological health. Case studies mention phobias, depression, and sexual disorders (3). Nonetheless, it has been ignored...
that female genital mutilation, representing a violation of someone’s physical intactness, can be classified as a psychological trauma according to DSM-IV and a potential cause of posttraumatic stress disorder (PTSD). Hence, the purpose of the study was to explore the relationship between female genital mutilation and psychiatric illnesses, especially PTSD. Despite the infliction of long-lasting somatic harm on most individuals, the fact that female genital mutilation is culturally embedded may represent a protective factor against the emergence of PTSD (4). In a subsidiary analysis, we also investigated whether or not the presence of PTSD is accompanied by declarative memory dysfunctions (deficits in active recall of previously encoded information), which are often found in other traumatized groups (5).

Method

The study was conducted over a 3-month period in the spring-time of 2003 in the capital of Senegal, Dakar. A group of 47 Senegalese women were interviewed in French by one of the authors (A.B.). Before the interviews, the researcher spent some time, usually about 2 days, establishing a solid relationship with the participants, creating a mutual sense of trust. The purposes of the study, as well as the normal ethical standards of anonymity, were explained to the women in detail.

The groups consisted of 23 circumcised and 24 uncircumcised women. They were recruited by a nurse, private contacts, and after the start of the survey, by former participants. Exclusion criteria were neurological disorders, psychotic disorders, current substance dependence, circumcision before the age of 5, and a lack of language skills (French). The participants’ ages ranged from 15 to 40 years, averaging 22.9 years (SD=4.2). In terms of education, the average was 11.5 years (SD=2.3). Twenty-one percent of the subjects were married, and 79% were single. More than 80% of the entire group had already faced traumatic life experiences. The most often named traumatic life event was the sudden death of a friend or a family member, which 66% of the participants had already witnessed. The two groups of circumcised and uncircumcised women did not differ significantly in terms of age, education, marital status, or traumatic life experiences. The age of circumcision ranged between 5 and 14 years; the mean was 8.2 years (SD=2.7). In none of the cases were local anesthetics or narcotics used.

A semistructured interview was performed comprising questions about demographic variables and questions concerning the event of circumcision and the subjects’ feelings toward it. The diagnostic interview in this study was the Mini International Neuropsychiatric Interview (6), a short structured interview that showed good reliability and validity compared to the Composite International Diagnostic Interview and to the Structured Clinical Interview for DSM-IV (7). Furthermore, a short form of the Traumatic Life Event Questionnaire (E.S. Kubany, unpublished, 1995) was used to assess other traumatic life experiences. Short-term and long-term memory were measured with the Rey Figure Test (8).

Results

All but one circumcised participant remembered the day of her circumcision as extremely appalling and traumatizing. Over 90% of the women described feelings of intense fear, helplessness, horror, and severe pain, and over 80% were still suffering from intrusive reexperiences of their circumcision. For 78% of the subjects, the event was performed unexpectedly and without any preliminary explanation. The psychiatric diagnoses based on the Mini International Neuropsychiatric Interview are presented in Table 1 for circumcised and uncircumcised women. Almost 80% of the circumcised women met criteria for affective or anxiety disorders. PTSD was particularly highly presented (30.4%). In the group of uncircumcised women, only one subject fulfilled the diagnostic criteria for an affective disorder.

In the copy trial of the Rey Figure Test, no differences among the three groups appeared. However, in both the immediate and delayed trials, participants with PTSD performed significantly worse than the circumcised women without PTSD and the uncircumcised women (immediate recall: F=5.31, df=2, 43, p<.01; with Bonferroni correction: participants with PTSD > circumcised women without PTSD (p<.05), uncircumcised women (p=.01); delayed recall: F=4.73, df=2, 43, p<.01; with Bonferroni correction: participants with PTSD > circumcised women without PTSD (p<.05), uncircumcised women (p=.01). Circumcised women without PTSD and uncircumcised women were indistinguishable in both parameters.

Discussion

The results of this study indicated that female genital mutilation is likely to cause various emotional disturbances, forging the way to psychiatric disorders, especially PTSD. The high rate of PTSD of more than 30% in this group is comparable to the rate of PTSD of early childhood abuse, which, on average, ranges between 30% and 50% (9). Despite the fact that female genital mutilation presents a part of the participants’ ethnic background, the results imply that cultural embedment does not protect against the development of PTSD and other psychiatric disorders. In agreement with other studies (5), PTSD rather than trauma was associated with declarative memory dysfunction.

Although the results of this study support the claim that female genital mutilation exerts a major negative impact on mental health beyond somatic complications, some caution is warranted in interpreting these results. The small group size presents an important limitation, and the results do not allow general conclusions about the prevalence of psychiatric disorders after female genital mutilation. The composition characteristics of the group (i.e.,

<table>
<thead>
<tr>
<th>Psychiatric Disorder</th>
<th>Circumcised Group (N=23)</th>
<th>Uncircumcised Group (N=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posttraumatic stress disorder</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Other anxiety disorders</td>
<td>6</td>
<td>26.2</td>
</tr>
<tr>
<td>Affective disorders</td>
<td>5</td>
<td>21.7</td>
</tr>
<tr>
<td>Other anxiety disorders</td>
<td>6</td>
<td>26.2</td>
</tr>
<tr>
<td>Affective disorders</td>
<td>5</td>
<td>21.7</td>
</tr>
<tr>
<td>No psychiatric diagnosis</td>
<td>5</td>
<td>21.7</td>
</tr>
</tbody>
</table>

TABLE 1. Psychiatric Disorders, Measured With the Mini International Neuropsychiatric Interview, French Version, in Circumcised and Uncircumcised Women
the low age and the high level of education) might also have an influence on the findings of this study. The attempt to interview women with lower education levels with a female interpreter did not work out because of the sensitive subject and the highly secret nature of female genital mutilation.

Moreover, the low prevalence of female genital mutilation in Senegal may be important for the occurrence of psychiatric disorders, possibly representing another risk factor for the establishment of PTSD. Since only 20% of the female population have undergone female genital mutilation, the circumcised women are often aware that their status is not generally accepted in society. Even more, in schools and other public facilities, people are increasingly taught about the negative effects, creating an atmosphere against female genital mutilation (2). As a consequence, the assimilation of the trauma may be different from other African countries where circumcision is more widespread. Of course, many other factors, such as coping style, other traumatic experiences, information before circumcision, and sexual experiences, may also play an important role in mental health outcome. Nevertheless, this study provides an additional strong argument for the provision of culturally grounded knowledge that can contribute to the eradication of female genital mutilation. In the spirit of the expanding role of psychiatrists and psychologists in documenting human rights abuses, it is of great importance to continue the research and to support women suffering from emotional difficulties. The alarmingly high rates of psychiatric disturbance among this group of circumcised women provide important evidence that researchers, as well as clinicians, have an obligation to focus more attention on the urgent needs of circumcised women.

Received March 18, 2004; revisions received April 13, 2004, and July 4, 2004; accepted July 28, 2004. From the University Hospital Hamburg-Eppendorf, Germany; and the Clinical Neuropsychology Unit, Hospital for Psychiatry and Psychotherapy. Address correspondence and reprint requests to Dr. Moritz, Clinical Neuropsychology Unit, Hospital for Psychiatry and Psychotherapy, Martinistraße 52, 20246 Hamburg, Germany; moritz@uke.uni-hamburg.de (e-mail).

The authors thank Rouguyatou Niang, Fatou Mata Hanne, Fousseni Nassakou, Prof. Dr. Naber, Dr. Frauke Teegen, Prof. Dr. Bernhard Dahme, Emily Page, Lena Jelinek, and Florentine Larbig.

References

Alcohol and Aldehyde Dehydrogenase Polymorphisms in Men With Type I and Type II Alcoholism

Young-Gyu Chai, Ph.D.
Dong-Yul Oh, M.D., Ph.D.
Eun Kee Chung, M.D., Ph.D.
Gil Sook Kim, M.D.
Leen Kim, M.D., Ph.D.
Yu-Sang Lee, M.D.
Ihn-Geun Choi, M.D., Ph.D.

Objective: The authors examined the genetic polymorphisms of alcohol dehydrogenase 2 and 3 (ADH2 and ADH3) and aldehyde dehydrogenase (ALDH2) in patients diagnosed as having Cloninger’s type I or type II alcoholism.

Method: Seventy-two alcoholic men and 38 nonalcoholic, healthy men were tested for the distribution of genotypes and alleles of ADH2, ADH3, and ALDH2. Forty-eight of the alcoholic men had type I alcoholism, and 24 had type II alcoholism.

Results: The frequencies of ADH2*1 and ADH3*2 alleles were significantly higher in men with type II alcoholism than in men with type I alcoholism and healthy men. The frequency of the ALDH2*1 allele was significantly higher in men with alcohol dependence than in healthy men.

Conclusions: The genetic characteristics of alcohol dehydrogenases in men with type I alcoholism were similar to those of healthy men, and the genetic characteristics of aldehyde dehydrogenase in men with type I alcoholism were similar to those of men with type II alcoholism. These findings suggest that the genetic characteristics of alcohol metabolism in type I alcoholism fall between nonalcoholism and type II alcoholism.

P

atients with alcohol dependence can be clinically clustered into heterogeneous groups. Cloninger (1) identified two genetic subtypes of alcoholics. Individuals with type I (milieu-limited) alcoholism have characteristics that include a later onset of alcohol-related problems, the development of psychological rather than physical dependence, and guilt related to their alcohol use. Subjects with type II (male-limited) alcoholism, on the other hand, manifest alcohol problems at an earlier age, show spontaneous or compulsive alcohol-seeking behavior, and are socially disruptive when drinking.

Although the pathogenesis of each clinical syndrome has not been clearly elucidated, it is likely that variation in alcohol elimination and acetaldehyde oxidation, genetically determined by alcohol dehydrogenases 1–3 (ADH1, ADH2, and ADH3) and aldehyde dehydrogenase (ALDH2), plays a central role (2). High blood concentration of acetaldehyde causes extremely unpleasant symptoms. Individuals possessing high-active ADH2*2 and ADH3*1 alleles generate acetaldehyde much more rapidly after alcohol consumption and are thus less tolerant of alcohol (3). The inactive ALDH2*2 allele, discovered only in Orientals, delays the clearance of acetaldehyde and protects against alcohol abuse (3).

To determine the genetic characteristics of alcohol metabolism in Cloninger’s type I and type II alcoholism, we analyzed the polymorphisms of ADH2, ADH3, and ALDH2 genes in healthy men and in men with type I or type II alcoholism.

Method

The subjects were 72 Korean male patients (mean age=45 years, SD=9) who met DSM-IV criteria for alcohol dependence and were classifiable definitely into type I (mean age at onset=36 years, SD=7) and type II (mean age at onset=22 years, SD=2) alcoholism according to the criteria of Cloninger (1). The patients had been admitted to Han-Gang Sacred Heart Hospital, Seoul National Mental Hospital, or Yong-In Mental Hospital. The comparison subjects were 38 unrelated healthy Korean men (mean age=34 years, SD=10) who visited the Health Screening Center in Han-Gang Sacred Heart Hospital. After a complete description of the study to the subjects, written informed consent was obtained.

The patients and comparison subjects were carefully assessed for their drinking patterns, alcohol-related problems, and psychiatric diagnoses by psychiatrists’ interview. The subjects had no major medical or psychiatric illnesses other than alcohol-related disorders in the patients.

Genomic DNA was extracted from peripheral blood leukocytes by using a standard phenol-chloroform method. We genotyped ADH2, ADH3, and ALDH2 by polymerase chain reaction amplification of DNA fragments containing targeted differences in base pairs. The positions and sequences of primers to amplify DNA fragments were based on minor modifications of the methods described by Xu et al. (4) for ADH1–3 and Crabb et al. (5) for ALDH2.

Genotype and allele frequency comparisons were performed by using the chi-square method with Monte Carlo simulation (6) as implemented by a computer program developed by Dave Curtis (http://www.mds.qmw.ac.uk/statgen/dcurtis/software.html).

Results

The frequencies of low-active ADH2 (ADH2*1/*1) and ADH3 (ADH3*1/*2 or ADH3*2/*2) genotypes were significantly higher in men with type II alcoholism (43% and 58%, respectively) than in men with type I alcoholism (23% and 24%, respectively). The allele and genotype fre-
quencies of ADH2 and ADH3 were not different between men with type I alcoholism and healthy men (Table 1).

The frequency of active ALDH2 (ALDH2*1/*1) genotype was significantly higher in men with alcohol dependence than in healthy men (96% versus 53%, respectively). Inactive ALDH2 (ALDH2*1/*2 or ALDH2*2/*2) genotype was present in only 4% of men with alcohol dependence, whereas 47% of the healthy men had inactive ALDH2. The frequencies of ALDH2 alleles and genotypes were not different between men with type I and type II alcoholism (Table 1).

The subjects in the overall study group, and in each comparison group, were in Hardy-Weinberg equilibrium.

Discussion

To our knowledge, this paper is the first report of a significant difference in allele and genotype frequencies of ADH2 and ADH3 between patients with type I and type II alcoholism. The allele and genotype frequencies of ADH2 and ADH3 were similar in men with type I alcoholism and healthy men. Healthy nonalcoholic men and men with type I alcoholism might retain high-active ADH2*1/*2, ADH2*2/*2, and ADH3*1/*1 genotypes that produce and cumulate acetaldehyde more rapidly, thus delaying or preventing the development of an alcohol problem.

The frequencies of ALDH2 alleles and genotypes were significantly different between men with alcohol dependence and healthy men. Healthy nonalcoholic men might retain an inactive ALDH2*1/*2 genotype that delays the clearance of acetaldehyde and protects against alcohol dependence.

The subjects included only a small number of Korean men, limiting generalizability of the findings. In addition, this is an association study, showing an association between certain genetic polymorphisms and diagnostic subgroups. This does not establish causality. The results suggest that the genotype that is expected to either metabolize ethanol to acetaldehyde more rapidly or slow the conversion of acetaldehyde to acetyl protects against an alcohol problem, mitigating the development of early-onset, severe alcohol dependence rather than decreasing the risk for developing alcohol dependence.

Subjects with type II alcoholism are genetically less similar to those of normal comparison subjects and genetic characteristics of aldehyde dehydrogenase similar to type I alcoholism might retain high-active ADH2*1/*2, ADH2*2/*2, and ADH3*1/*1 genotypes that produce and cumulate acetaldehyde more rapidly, thus delaying or preventing the development of an alcohol problem.

The results of our study suggest that type I alcoholics had genetic characteristics of alcohol dehydrogenases similar to those of normal comparison subjects and genetic characteristics of aldehyde dehydrogenase similar to those of type II alcoholics. The ALDH2 gene may play a

<table>
<thead>
<tr>
<th>Groupa</th>
<th>Allele Frequency</th>
<th>Genotype Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADH2*1b</td>
<td>ADH2*2b</td>
</tr>
<tr>
<td>Subjects with type I alcoholism (N=48)</td>
<td>0.52 0.48</td>
<td>0.23 0.58 0.19</td>
</tr>
<tr>
<td>Subjects with type II alcoholism (N=23)</td>
<td>0.72 0.28</td>
<td>0.43 0.57 0.00</td>
</tr>
<tr>
<td>Healthy subjects (N=38)</td>
<td>0.47 0.53</td>
<td>0.18 0.58 0.24</td>
</tr>
<tr>
<td></td>
<td>ADH3*1d</td>
<td>ADH3*2d</td>
</tr>
<tr>
<td>Subjects with type I alcoholism (N=46)</td>
<td>0.88 0.12</td>
<td>0.76 0.24 0.00</td>
</tr>
<tr>
<td>Subjects with type II alcoholism (N=19)</td>
<td>0.68 0.32</td>
<td>0.42 0.53 0.05</td>
</tr>
<tr>
<td>Healthy subjects (N=35)</td>
<td>0.89 0.11</td>
<td>0.77 0.23 0.00</td>
</tr>
<tr>
<td></td>
<td>ADH3*1e</td>
<td>ADH3*2e</td>
</tr>
<tr>
<td>Subjects with type I alcoholism (N=48)</td>
<td>0.98 0.02</td>
<td>0.96 0.04 0.00</td>
</tr>
<tr>
<td>Subjects with type II alcoholism (N=23)</td>
<td>0.98 0.02</td>
<td>0.96 0.04 0.00</td>
</tr>
<tr>
<td>Healthy subjects (N=38)</td>
<td>0.76 0.24</td>
<td>0.53 0.47 0.00</td>
</tr>
</tbody>
</table>

a Numbers of individuals in each group are given. Some individuals were excluded because some exons did not amplify well or gave ambiguous results.
b Comparison of subjects with type II alcoholism with subjects with type I alcoholism (χ²=6.56, p=0.01, SE=0.003) and healthy subjects (χ²=7.85, p=0.005, SE=0.001).
c Comparison of subjects with type II alcoholism with subjects with type I alcoholism (χ²=7.92, p=0.02, SE=0.005) and healthy subjects (χ²=9.53, p=0.009, SE=0.000).
d Comparison of subjects with type II alcoholism with subjects with type I alcoholism (χ²=6.02, p=0.02, SE=0.001) and healthy subjects (χ²=7.13, p=0.008, SE=0.002).
e Comparison of subjects with type II alcoholism with subjects with type I alcoholism (χ²=7.00, p=0.030, SE=0.004) and healthy subjects (χ²=7.77, p=0.021, SE=0.002).
f Comparison of healthy subjects with subjects with type I alcoholism (χ²=19.71, p=0.000, SE=0.000) and subjects with type II alcoholism (χ²=10.08, p=0.002, SE=0.000).
g Comparison of healthy subjects with subjects with type I alcoholism (χ²=22.65, p=0.000, SE=0.000) and subjects with type II alcoholism (χ²=12.36, p=0.002, SD=0.000).

http://ajp.psychiatryonline.org

Am J Psychiatry 162:5, May 2005
major role in developing alcohol dependence, and both ADH2 and ADH3 genes may play the role in severity of alcohol dependence. We presume that the genetic characteristics of alcohol metabolism of type I alcoholism fall between nonalcoholism and type II alcoholism.

References

Association of Genetic Variants in Alcohol Dehydrogenase 4 With Alcohol Dependence in Brazilian Patients

Camila Guindalini, B.Sc.
Sandra Scivoletto, Ph.D.
Ricardo G.M. Ferreira, B.Sc.
Gerome Breen, Ph.D.
Monica Zilberman, Ph.D.
Marco Aurélio Peluso, Ph.D.
Mayana Zatz, Ph.D.

Objective: The authors evaluated the association of three functional promoter polymorphisms of the ADH4 gene with alcohol dependence.

Method: DNA from 92 alcohol-dependent patients and 92 healthy comparison subjects was genotyped for all three polymorphisms.

Results: Variants at the –75 base-pair (bp) (C allele) and –159 bp (A allele) positions were associated with alcohol dependence. Individuals with haplotypes carrying both risk alleles showed an odds ratio of 2.9.

Conclusions: These preliminary results suggest that ADH4 may play a role in the etiology of alcohol dependence. The association requires further study and replication but is functionally plausible and has a large effect size.

Alcohol metabolism strongly influences drinking behavior and the likelihood of being dependent (1). Most ethanol is oxidized to acetaldehyde and acetate, mainly by alcohol dehydrogenases (ADH types 1–4) and aldehyde dehydrogenase (ALDH) (2). Functional allelic variants of ADH1B (*2) and ADH1C (*1) giving ADHs with higher maximum velocity of an enzyme-catalyzed reaction for a given value of enzyme concentration reduce the risk for alcohol dependence (3–5).

The gene for ADH4 (ADH4/*ADH) is mapped to 4q22 (6). Human ADH4 enzyme is found mainly in the liver, and at intoxicating levels of alcohol it may account for as much as 40% of the total ethanol oxidation rate (7). Edenberg et al. (8) reported three polymorphisms in the promoter of ADH4: –192 base-pair (bp) (T/A), –159 bp (G/A), and –75 bp (A/C). These authors analyzed the effects of four different haplotypes (T,A,G; T,G,A; A,G,C; and T,A,A) on gene expression. Two promoters with A at –75 bp had higher ac-
Association of Genetic Variants in Alcohol Dehydrogenase 4 With Alcohol Dependence in Brazilian Patients

Camila Guindalini, B.Sc.
Sandra Scivoletto, Ph.D.
Ricardo G.M. Ferreira, B.Sc.
Gerome Breen, Ph.D.
Monica Zilberman, Ph.D.
Marco Aurélio Peluso, Ph.D.
Mayana Zatz, Ph.D.

Objective: The authors evaluated the association of three functional promoter polymorphisms of the ADH4 gene with alcohol dependence.

Method: DNA from 92 alcohol-dependent patients and 92 healthy comparison subjects was genotyped for all three polymorphisms.

Results: Variants at the –75 base-pair (bp) (C allele) and –159 bp (A allele) positions were associated with alcohol dependence. Individuals with haplotypes carrying both risk alleles showed an odds ratio of 2.9.

Conclusions: These preliminary results suggest that ADH4 may play a role in the etiology of alcohol dependence. The association requires further study and replication but is functionally plausible and has a large effect size.

References


Alcohol metabolism strongly influences drinking behavior and the likelihood of being dependent (1). Most ethanol is oxidized to acetaldehyde and acetate, mainly by alcohol dehydrogenases (ADH types 1–4) and aldehyde dehydrogenase (ALDH) (2). Functional allelic variants of ADH1B (∗2) and ADH1C (∗1) giving ADHs with higher maximum velocity of an enzyme-catalyzed reaction for a given value of enzyme concentration reduce the risk for alcohol dependence (3–5).

The gene for ADH4 (ADH4/*ADH) is mapped to 4q22 (6). Human ADH4 enzyme is found mainly in the liver, and at intoxicating levels of alcohol it may account for as much as 40% of the total ethanol oxidation rate (7). Edenberg et al. (8) reported three polymorphisms in the promoter of ADH4: –192 base-pair (bp) (T/A), –159 bp (A allele) and –75 bp (A/C). These authors analyzed the effects of four different haplotypes (T/A; T/G; A/G; and T/T) on gene expression. Two promoters with A at –75 bp had higher ac-
Alcohol-Dependent Patients and Nonalcoholic Comparison Subjects

Method

 Ninety-two patients with ICD-10-defined alcohol dependence were ascertained at the Institute of Psychiatry, Sao Paulo. Fifty-one of the patients were men and 41 were women; their mean age was 47.3 years (SD=9.7). The 92 comparison subjects were enrolled at the Human Genome Research Center, University of Sao Paulo. Fifty-one of the comparison subjects were men, and 41 were women; their mean age was 45.3 years (SD=11.7). All comparison subjects were screened with CAGE questionnaire (9) for potential alcohol-related disorders. The recommended cutoff score on this instrument for alcohol problems is 2, but we adopted a score of 1 for the selection of the comparison subjects. The healthy subjects were matched for ethnic background with the patients: 76% were Caucasian, and 24% were African Brazilian. All subjects gave written informed consent.

Genomic DNA was extracted from whole blood according to standard procedures (10) and submitted to polymerase chain reaction (PCR) amplification (8). The PCR products were sequenced using ABI-BigDye kit on an Applied Biosystems Model 377 (Applied Biosystems, Foster City, Calif.). Fifty healthy individuals were typed for ADH1B polymorphism as described elsewhere (3).

We used both chi-square tests and Fisher’s exact tests to determine the significance of differences between groups. GENECOUNTING software (version 1.3, March 2003) (11) was used to estimate the haplotype frequencies and to evaluate pairwise measures of linkage disequilibrium with Cramer’s V statistic. Corresponding odds ratios were calculated by using the method of Woolf (12).

Results

Genotype frequencies for the ADH4 gene among alcohol-dependent patients and comparison subjects were significantly different for the −75 bp and −159 bp polymorphisms. The differences were still significant after taking ethnicity into account (Table 1) and after logistic regression analyses controlling for ethnicity were carried out (data not shown). No association was observed for the −192 bp polymorphism (p>0.05). The homogeneity of the study group for the ADH4 polymorphisms allows the grouping of the data from the two ethnic groups. The −75C allele (odds ratio=1.6, 95% CI=1.05–2.4, p<0.05) and the −159A allele (odds ratio=2.2, 95% CI=1.3–3.3, p<0.001) were associated with alcohol dependence (data not shown). The GENECOUNTING program generated eight different haplotypes among patients and comparison subjects (Table 2). The haplotypes A,A,C and T,A,C were significantly more frequent among patients (p=0.01 and p=0.02, respectively) and were considered risk haplotypes. Among comparison subjects the haplotypes T,A,A and A,G,C were overrepresented (p=0.01 and p=0.02, respectively) and were considered protective.

Haplotypes with the allele −75C and the allele −75A were then compared. The same analysis was done with the −159A and −159G alleles and finally for haplotypes with both −75C and −159A alleles versus all others. The combination of both risk alleles on one haplotype showed an odds ratio of 2.9 (95% CI=4.73–1.89, p<0.00001).

To test linkage disequilibrium between ADH1B functional polymorphism and ADH4 single nucleotide polymorphisms (SNPs), 50 healthy individuals were genotyped. No significant linkage disequilibrium was observed among them (Cramer’s V=0.095 for −75 bp, 0.136 for −159 bp).

---

**Table 1. Genotype Distribution of the Polymorphisms at −75 Base-Pair (bp), −159 bp, and −192 bp Positions Among Alcohol-Dependent Patients and Nonalcoholic Comparison Subjects**

<table>
<thead>
<tr>
<th>Group</th>
<th>−75 bp</th>
<th>−159 bp</th>
<th>−192 bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>All subjects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol-dependent patients (N=92)</td>
<td>5</td>
<td>78</td>
<td>9</td>
</tr>
<tr>
<td>Comparison subjects (N=92)</td>
<td>21</td>
<td>67</td>
<td>4</td>
</tr>
<tr>
<td>Caucasian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol-dependent patients (N=70)</td>
<td>4</td>
<td>57</td>
<td>9</td>
</tr>
<tr>
<td>Comparison subjects (N=70)</td>
<td>14</td>
<td>53</td>
<td>3</td>
</tr>
<tr>
<td>African Brazilian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol-dependent patients (N=22)</td>
<td>1</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>Comparison subjects (N=22)</td>
<td>7</td>
<td>14</td>
<td>1</td>
</tr>
</tbody>
</table>

a Significance of difference between alcohol-dependent patients and comparison subjects determined by chi-square test.

**Table 2. Distribution of Eight Haplotypes Generated by the GENECOUNTING Program Among Alcohol-Dependent Patients and Nonalcoholic Comparison Subjects**

<table>
<thead>
<tr>
<th>Percentage of the Subjects With Haplotype</th>
<th>Number of Subjects by Genotype</th>
<th>p²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haplotype</td>
<td>patients</td>
<td></td>
</tr>
<tr>
<td>A,A,A</td>
<td>5.55</td>
<td>0.00</td>
</tr>
<tr>
<td>T,A,A</td>
<td>2.56</td>
<td>8.31</td>
</tr>
<tr>
<td>A,G,A</td>
<td>0.00</td>
<td>1.35</td>
</tr>
<tr>
<td>T,G,A</td>
<td>44.71</td>
<td>49.98</td>
</tr>
<tr>
<td>A,A,C</td>
<td>22.82</td>
<td>11.64</td>
</tr>
<tr>
<td>T,A,C</td>
<td>21.88</td>
<td>9.63</td>
</tr>
<tr>
<td>A,G,C</td>
<td>6.51</td>
<td>17.12</td>
</tr>
<tr>
<td>T,G,C</td>
<td>0.96</td>
<td>1.98</td>
</tr>
</tbody>
</table>

a Significance of difference between alcohol-dependent patients and comparison subjects determined by chi-square test confirmed by simulations.
bp, and 0.062 for −192 bp polymorphisms) (p>0.05) (data not shown).

Discussion

Two ADH4 polymorphisms and their haplotypes were associated with greater risk of developing alcohol dependence. In our study group, individuals who carried a cytosine at −75 bp of the ADH4 gene and adenine at −159 bp were three times more likely to develop alcohol dependence. All other probable haplotypes showed a protective or neutral association. We investigated the possibility that the association of these ADH4 polymorphisms with alcohol dependence could be attributable to linkage disequilibrium with the ADH1B locus. This seems not to be the case, given the low linkage disequilibrium between the SNPs. Additionally, published data have shown that there is no significant linkage disequilibrium between the ADH1 and ADH4 loci (13).

Edenberg et al. (8) reported a higher expression for the −75A allele and suggested it should be protective against alcohol dependence. Moreover, they reported that the T,A,C and T,A,A promoters had lower and higher levels of expression, respectively. We agree with these authors because T,A,C—found mainly among patients—was a risk factor in our study and T,A,A—significantly overrepresented in the comparison group—was protective. The effect of our other risk haplotype—A,A,C—on promoter function was not measured by Edenberg et al. The only exception was that the haplotype A,G,C, reported by them to be a risk factor in our study and T,A,A promoted in the comparison group, was apparently protective in our sample.

In short, our results suggest a role for common ADH4 promoter polymorphisms in the etiology of alcoholism. Given that our study group is small, and there is a possibility of genetic stratification, these findings need to be replicated in larger and different populations. In addition, further genotyping of the other ADH gene polymorphisms will be necessary in large samples to investigate the possibility of linkage disequilibrium. However, this association is with common alleles/haplotypes and is functionally plausible.

References


Received June 20, 2003; revisions received March 12 and July 1, 2004; accepted July 28, 2004. From the Human Genome Research Center, Department of Biology, Institute of Biosciences, and GREA, Interdisciplinary Group of Studies on Alcohol and Drugs, Institute and Department of Psychiatry, Medicine Faculty; and Department of Parasitology, Institute of Biomedical Sciences, University of Sao Paulo, Brazil; and the Section of Genetics, Division of Psychological Medicine and SGDP Research Centre, Institute of Psychiatry, Kings College London, De Crespigny Park, London. Address correspondence and reprint requests to Dr. Zatz, Human Genome Research Center, Department of Biology, Institute of Biosciences, University of Sao Paulo, Rua do Matão, Travessa 13, Number 106, Cidade Universitária CEP: 05508-090, Sao Paulo SP, Brazil; mayazatz@usp.br (e-mail).

Supported in part by the Fundação de Amparo a Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico.

The authors thank Constâncio Urbani, Patrícia Hochgraf, Antonia Cerqueira, and Elisangela Quedas.
Substance Abuse in First-Episode Bipolar I Disorder: Indications for Early Intervention

Christopher Baethge, M.D.
Ross J. Baldessarini, M.D.
Hari-Mandir Kaur Khalsa, A.B.
John Hennen, Ph.D.
Paola Salvatore, M.D.
Mauricio Tohen, M.D., Dr.P.H.

Objective: This study clarified the early characteristics of substance use disorders in patients with first-episode bipolar I disorder.

Method: The authors evaluated substance use disorders, associated factors, and clinical course, prospectively, in the first 2 years of DSM-IV bipolar I disorder with standardized methods.

Results: Baseline substance use disorder was found in 33% (37 of 112) of the patients at baseline and in 39% at 24 months. Anxiety disorders were more frequent in the patients with than without substance use disorder (30% and 13%, respectively). Associations of alcohol dependence with depressive symptoms and cannabis dependence with manic symptoms were suggested. Patients using two or more substances had worse outcomes.

Conclusions: Since substance use disorders were frequent from the beginning of bipolar I disorder and were associated with anxiety disorders and poor outcome, early interventions for substance use disorder and anxiety might improve later outcome.

(Am J Psychiatry 2005; 162:1008–1010)
Of 22 subjects using one substance only (group 2), 17 used alcohol, four cannabis, and one heroin; the 15 in group 3 abused two or more substances (Table 1). Dependence was ranked as follows: 23 with alcohol, 11 cannabis, five cocaine, and one hallucinogens; abuse was ranked as follows: seven with alcohol, four cannabis, one cocaine, and two hallucinogens. For both dependence and abuse, there was considerable overlap in use of different drugs (average=1.8 per substance use disorder case). The prevalence of substance use disorder rose to 38.8% at 24 months (31 of 80 with a second SCID assessment).

Polydrug use subjects (group 3) versus nonabusers (group 1) were younger, less well educated, more likely to have a family history of psychiatric illness, and to be seen initially in mixed states. Subgroups with no abuse and monodrug use were similar in all of these baseline measures, and the subgroups did not differ by sex or race. Co-morbidity with any anxiety disorder was more common with substance use disorders. Patients with polydrug use spent more time ill during follow-up (especially in mixed episodes). Syndromal recovery was similar across subgroups. Other morbidity at 2 years (including initial depression severity and time spent in major depression/dysthymia or mania) did not differ significantly among subgroups (Table 1).

We also compared 2-year outcomes of patients with no substance use disorder versus alcohol-dependent (N=23) versus cannabis-dependent patients (N=11). Patients with no substance use disorder spent similar proportions of time in manic and depressive illness (mean=14.4%, SD=19.5%, versus mean=13.5%, SD=21.9%; ratio=1:1.1). Cannabis-dependent subjects spent more time in mania (mean=26.8%, SD=34.0%, versus mean=11.5%, SD=19.1%; ratio=2.3:1), whereas alcohol-dependent patients spent much more time depressed (mean=11.9%, SD=22.9%, in mania versus mean=21.9%, SD=29.5%, in depression; ratio=1:1.8). However, these findings did not reach statistical significance, and there was a considerable overlap in the use of both drugs.

**Discussion**

This preliminary assessment indicated that the prevalence of substance use disorder was already substantial among first-episode, DSM-IV bipolar I subjects (33%), reaching 39% by 24 months. Syndromal recovery from index mania occurred at similar rates across substance use disorder subgroups. However, overall morbidity during follow-up was particularly severe among patients with polysubstance use disorder, whereas patients with monosubstance use disorder had surprisingly good early outcomes. A previously suggested (3) association of alcohol with depression and cannabis with mania was found and warrants further study.

Of note, patients with comorbid substance use disorder were more likely than nonusers to be diagnosed with a DSM-IV anxiety disorder. A link between anxiety and substance use disorders in major affective disorders has been proposed (4), but the cause-effect relationships remain unclear. Substance use might sometimes ameliorate anxi-
ety symptoms or anxiety worsened by substance use disorder. The evident association suggests the testable hypothesis that very early clinical interventions in new bipolar I disorder aimed at limiting both anxiety symptoms and substance use disorder might contribute to improved outcomes.

Presented in part at the 157th annual meeting of the American Psychiatric Association, New York, May 1–6, 2004. Received Jan. 27, 2004; revisions received April 30 and May 28, 2004; accepted June 14, 2004. From the Department of Psychiatry and the Neuroscience Program, Harvard Medical School, Boston; the Bipolar & Psychotic Disorders Program and International Consortium for Bipolar Disorder Research, Mailman Research Center, McLean Hospital; the Department of Clinical Neuroscience, University of Parma, Italy; and Lilly Research Laboratories, Indianapolis. Address correspondence and reprint requests to Dr. Baethge, Department of Psychiatry, Harvard Medical School, McLean Hospital, 115 Mill St., Belmont, MA 02478-9106; cbaethge@mclean.harvard.edu (e-mail).

Supported by the Max Kade Foundation, a National Alliance for Research on Schizophrenia and Depression Young Investigator Award (to Dr. Baethge), the Bruce J. Anderson Foundation, and the McLean Private Donors Research Fund (to Dr. Baldessarini).

References

Brief Report

Predominant Role of the 9-Hydroxy Metabolite of Risperidone in Elevating Blood Prolactin Levels

Rikus Kneegtering, M.D., Ph.D.
Pepijn Baselmans, M.D.
Stynke Castelein, M.Sc.
Fokko Bosker, M.Sc., Ph.D.
Richard Bruggeman, M.D., Ph.D.
Robert J. van den Bosch, M.D., Ph.D.

Objective: The atypical antipsychotic risperidone significantly raises plasma prolactin levels in patients, but clozapine, olanzapine, and quetiapine do not. The differences in neuroendocrine response may be connected with the metabolism of the medications. The authors examined the contributory role of risperidone’s active metabolite 9-hydroxy-risperidone by measuring plasma concentrations of risperidone, 9-hydroxy-risperidone, and prolactin.

Method: Blood samples taken from 25 patients with psychotic disorders following 6 weeks of treatment with risperidone (mean dose = 3 mg/day) were examined. Mean plasma concentrations of risperidone, 9-hydroxy-risperidone, and prolactin were 4.6, 19.4, and 49.3 ng/ml, respectively.

Results: The oral dose of risperidone correlated significantly with plasma concentrations of risperidone, 9-hydroxy-risperidone, and prolactin. The plasma concentration of 9-hydroxy-risperidone, but not of risperidone, correlated significantly with increases in plasma prolactin.

Conclusions: These data suggest that the 9-hydroxy metabolite plays a predominant role in risperidone’s effect on prolactin release.

(Tam J Psychiatry 2005; 162:1010–1012)
BRIEF REPORTS

Predominant Role of the 9-Hydroxy Metabolite of Risperidone in Elevating Blood Prolactin Levels

Rikus Knegethering, M.D., Ph.D.
Pepijn Baselmans, M.D.
Stynke Castelein, M.Sc.
Fokko Bosker, M.Sc., Ph.D.
Richard Bruggeman, M.D., Ph.D.
Robert J. van den Bosch, M.D., Ph.D.

Objective: The atypical antipsychotic risperidone significantly raises plasma prolactin levels in patients, but clozapine, olanzapine, and quetiapine do not. The differences in neuroendocrine response may be connected with the metabolism of the medications. The authors examined the contributory role of risperidone’s active metabolite 9-hydroxy-risperidone by measuring plasma concentrations of risperidone, 9-hydroxy-risperidone, and prolactin.

Method: Blood samples taken from 25 patients with psychotic disorders following 6 weeks of treatment with risperidone (mean dose=3 mg/day) were examined. Mean plasma concentrations of risperidone, 9-hydroxy-risperidone, and prolactin were 4.6, 19.4, and 49.3 ng/ml, respectively.

Results: The oral dose of risperidone correlated significantly with plasma concentrations of risperidone, 9-hydroxy-risperidone, and prolactin. The plasma concentration of 9-hydroxy-risperidone, but not of risperidone, correlated significantly with increases in plasma prolactin.

Conclusions: These data suggest that the 9-hydroxy metabolite plays a predominant role in risperidone’s effect on prolactin release.

References

The increase in prolactin levels associated with the use of antipsychotics is caused by the blockage of dopamine D2 receptors in the anterior pituitary (1), which is located outside the blood-brain barrier. Most classical antipsychotics are potent D2 receptor antagonists, and, therefore, they induce hyperprolactinemia. Atypical antipsychotic medications are generally less potent D2 receptor antagonists, which is in line with the mild and transient increases in prolactin that have been observed in patients. Risperidone has a moderate potency (2, 3), and yet it produces
profound and long-lasting prolactin elevation (4). To date, neither antipsychotics nor their metabolites are known to have any direct effects on serum prolactin levels.

Kapur et al. (1, 5) proposed that, compared with olanzapine and quetiapine, risperidone's poor penetration of the blood-brain barrier might explain the drug's greater tendency to cause hyperprolactinemia (1, 5). Risperidone's metabolite, 9-hydroxy-risperidone, may be a confounding factor. Risperidone is metabolized by CYP2D6 and, to a lesser extent, by CYP3A4 into 9-hydroxy-risperidone. In contrast to the metabolites of most other antipsychotics, 9-hydroxy-risperidone contributes significantly to the therapeutic effect of risperidone (6, 7). We are not aware of any clinical studies that have examined the role of the 9-hydroxy metabolite in the side effects of risperidone. To test the hypothesis that this metabolite contributes considerably to the drug's tendency to elevate prolactin levels, we measured the plasma levels of risperidone, 9-hydroxy-risperidone, and prolactin in patients who had been treated with risperidone for 6 weeks.

**Method**

Only patients being treated with risperidone as monotherapy were eligible for inclusion in the study. Blood samples were collected between 8.30 and 9.00 a.m. After liquid-liquid extraction, plasma risperidone and its 9-hydroxy metabolite were measured by reversed-phase high-performance liquid chromatography with ultraviolet detection. Detection limits of risperidone and the 9-hydroxy metabolite were 2 and 1 ng/ml, respectively. The MAIA clone kit (Ares Serono Diagnostics, Milan, Italy), which enables measurement of prolactin concentrations over a range of 6.0–10,000 ME/liter, was used to measure plasma prolactin levels.

The plasma concentration measurements were normalized by converting them to their natural logarithm. Pearson correlations were calculated, with the level of significance set at $p < 0.05$.

**Results**

The study group consisted of 25 patients (14 male, 11 female; mean age=25 years, SD=7, range=17–52) who had psychotic disorders diagnosed according to DSM-IV criteria (20 patients had diagnoses within the schizophrenia spectrum and five had other psychotic disorders). The mean dose of risperidone was 3 mg/day (range=1–6).

Mean plasma concentrations of risperidone, 9-hydroxy-risperidone, and prolactin were 4.6 ng/ml (SD=1.5), 19.4 ng/ml (SD=3.3), and 49.3 ng/ml (SD=7.2), respectively. Oral risperidone dose was significantly correlated with plasma risperidone levels ($r=0.52$, $N=25$, $p=0.008$), as well as plasma 9-hydroxy-risperidone levels ($r=0.57$, $N=25$, $p=0.003$) and plasma prolactin levels ($r=0.45$, $N=25$, $p<0.03$). In contrast to plasma risperidone, only 9-hydroxy-risperidone was significantly correlated with plasma prolactin levels ($r=0.52$, $N=25$, $p=0.008$). No significant correlations were found between plasma risperidone levels and plasma prolactin levels ($r=0.23$, $N=25$, $p=0.26$) or between plasma risperidone levels and 9-hydroxy-risperidone levels ($r=0.29$, $N=25$, $p=0.16$). Figure 1 shows a scatterplot of the correlations between serum risperidone and 9-hydroxy-risperidone compared with serum prolactin.

**Discussion**

The degree of hyperprolactinemia observed after 6 weeks of treatment with risperidone was in accordance with previous findings, which have shown prolactin levels exceeding those observed with most classical and atypical antipsychotics (1, 4, 5). Prolactin levels were correlated with levels of the 9-hydroxy metabolite, but not with those of risperidone.

Risperidone is less lipophilic than most antipsychotics, which is reflected by its lower brain-to-plasma ratio (6, 7). The relatively higher peripheral concentration of risperidone might explain the marked transient increase of prolactin, but not the sustained elevation on which it is superimposed (8). To understand this sustained prolactin elevation, the following factors should be considered. Although the potency of risperidone to stimulate prolactin release is comparable to that of 9-hydroxy-risperidone (9), the drug and the metabolite have different pharmacokinetics. The value of 9-hydroxy-risperidone for protein binding (77.4% versus 90.0%) (10) and its partition coefficient (log P: 2.32 versus 3.04) (10) both indicate that the metabolite is less lipophilic than risperidone. Moreover, the half-life of 9-hydroxy-risperidone (20 hours) is much longer than that of risperidone (2–4 hours) (11).
Compared with risperidone, the active metabolite has a similar dopamine D\textsubscript{2} receptor affinity, a lower brain-to-plasma ratio, and a longer half-life; these data suggest that 9-hydroxy-risperidone played a predominant role in the strong prolactin elevation that was observed during treatment with risperidone in our study. Arguably, these data also explain risperidone's high propensity to cause amenorrhea, galactorrhea, and sexual side effects (12).

Received Feb. 2, 2004; revision received April 30, 2004; accepted June 14, 2004. From the Department of Psychiatry, University Hospital Groningen, Groningen, the Netherlands. Address correspondence and reprint requests to Dr. Knegtering, Hanzeplein 1, PO Box 30.001, 9700 RB Groningen, the Netherlands; H.Knegtering@psy.umcg.nl or rikusknegtering@home.nl (e-mail).

References

Brief Report

The Effects of Clozapine and Risperidone on Spatial Working Memory in Schizophrenia

Susan R. McGurk, Ph.D.
Cameron Carter, M.D.
Robert Goldman, Ph.D.
Michael F. Green, Ph.D.
Stephen R. Marder, M.D.
Haiyi Xie, Ph.D.
Nina R. Schooler, Ph.D.
John M. Kane, M.D.

Objective: The purpose of this investigation was to evaluate the effects of clozapine and risperidone on spatial working memory in patients with schizophrenia.

Method: Spatial working memory performance was evaluated at baseline and after 17 and 29 weeks in 97 patients with schizophrenia participating in a multisite trial.

Results: Compared with baseline performance while receiving conventional antipsychotic medication, risperidone improved, and clozapine worsened, spatial working memory performance.

Conclusions: The differential effects of these medications on spatial working memory may be due to the anticholinergic effects of clozapine and prefrontal dopamine-enhancing effects of risperidone.

Spatial working memory is a neuropsychological function of special interest in schizophrenia. First, nonhuman primate and human studies have demonstrated that spatial working memory processes rely on intact functioning of the dorsolateral prefrontal cortex (1, 2), an area associated with core deficits in schizophrenia. Second, studies that have evaluated spatial working memory using a variety of spatial delayed-response paradigms have demonstrated impairments in patients with schizophrenia (3, 4). Third, persons at risk for the illness demonstrate impairments in spatial working memory, including first-degree biological relatives (5) and high-risk adolescents who later develop schizophrenia (6), indicating that spatial working memory may be an endophenotype for schizophrenia. Fourth, spatial working memory deficits are associated with impaired community functioning, including work (7). Thus, spatial working memory taps neural processes that are of fundamental interest in schizophrenia, and impairments may indicate risk for illness and impaired community function.

Spatial working memory performance is also of interest in the evaluation of medications because it is altered by drugs commonly used in the treatment of schizophrenia. For example, in a 4-week, double-blind, randomized trial, risperidone improved, while haloperidol worsened, spatial working memory performance, an effect largely due to coadministration of benzotropine with haloperidol (8). Because clozapine exhibits potent antimuscarinic activity in vitro at a magnitude similar to that seen with benzotropine (9, 10), it may exert a deleterious effect on spatial working memory. Such an effect is suggested by impairment in visual memory (11, 12) and verbal working memory (13) during clozapine treatment. A more recent controlled trial comparing clozapine, risperidone, olanzapine, and haloperidol (14) indicated that verbal working memory and visual memory were unchanged from baseline in patients receiving clozapine and risperidone. Spatial working memory was not evaluated.

In addition, nonhuman primate studies have demonstrated that spatial working memory is impaired by D1 antagonists and improved by D1 agonists (15), indicating that spatial working memory is modulated, in part, by prefrontal dopaminergic processes. Thus, beneficial effects of risperidone compared with haloperidol on working memory (16) may be due to both its relative lack of anticholinergic properties (8, 17) as well as its ability to enhance dopamine activity in the prefrontal cortex (17).

The current study evaluated the effects of clozapine and risperidone on performance of a computerized assessment of spatial working memory in patients with moderately refractory schizophrenia or schizoaffective disorder. It was hypothesized that relative to clozapine, risperidone would improve spatial working memory.

Method

This study was part of a larger 6-month, multicenter, double-blind, randomized controlled trial comparing clozapine (target dose: 500 mg/day) and risperidone (target dose: 6 mg/day) in patients with moderately treatment-resistant schizophrenia. Clinical and adverse event data have been reported elsewhere (unpublished study of N.R. Schooler et al.). Subjects treated with clozapine were less likely to discontinue treatment for lack of efficacy (15%) than those treated with risperidone (38%) and showed more improvement globally and in asociality. Other reasons for discontinuation included withdrawal of consent and treatment-related side effects. However, the proportion of subjects meeting an a priori criterion of psychosis symptom improvement (20% improvement on at least one of four psychotic symptoms) did not significantly differ between the two groups (risperidone: 57%; clozapine: 71%). A previous report of this study indicated no medication ef-
-effects on social competence or problem solving (18). All subjects provided informed written consent for their participation.

Subjects

Study participants met DSM-IV criteria for schizophrenia or schizoaffective disorder as determined by a diagnostic checklist based on a structured interview. Subjects’ illness had shown evidence of treatment resistance, defined as at least one trial of a conventional antipsychotic at a dose equivalent to 600 mg/day of chlorpromazine, a second trial of a different conventional antipsychotic at a dose equivalent to 250–500 mg/day of chlorpromazine, and at least a moderate severity score on one of the Brief Psychiatric Rating Scale (BPRS) (19) psychotic symptom subscale items or on one of the Scale for Assessment of Negative Symptoms (20) global subscales. Last, subjects were 18–60 years of age and were living in the community or judged potentially dischargeable from the hospital. Exclusion criteria were diagnosis of neuroleptic malignant syndrome with recurrence upon rechallenge, history of CNS pathology, pregnancy, or mental retardation that precluded understanding study participation or assessment procedures.

Of the subjects recruited for the parent study (N=107), 97 completed baseline spatial working memory tests. Fifty subjects were randomly assigned to treatment with risperidone, and 47 received clozapine. Demographic variables (age, gender, education, ethnicity) by drug group are presented in Table 1.

Assessments

The BPRS was used to assess clinical symptoms. Composite scores of the psychotic symptom subscale (hallucinations, delusions, suspiciousness, and conceptual disorganization) and the negative symptoms anergia subscale (emotional withdrawal, blunted affect, motor retardation) were used in statistical analyses. Spatial working memory was measured using a computerized delayed response test (8) imposing one of two delays (5 or 15 seconds) between stimulus presentation (a black circle occurring in one of eight possible locations) and response (subjects pointed to the location of the target stimulus at the end of the delay). Thirty-two trials of each delay condition were presented. The dependent measure was the number of correctly identified targets at each delay. The test was approximately 20 minutes in duration. Practice trials, which were not included in the total score, consisted of 16 target presentations. For the initial eight practice trials, the target remained on the screen for 5 seconds and overlapped with the eight target choice displays. Because there was no delay between target presentation and response, accurate responding did not require memory, but required the ability to perceive test stimuli and point to the target. The next eight practice trials were identical to the 5-second delay trials and were also administered to ensure that subjects understood test procedures.

Procedure

Clinical and cognitive assessments were performed at baseline and 17 and 29 weeks after initiation of clozapine or risperidone treatment. Clozapine was initiated at 12.5 mg/day and gradually titrated to 500 mg/day by day 28. Risperidone was initiated at 1 mg/day and gradually titrated to 6 mg/day by day 15. Simultaneous with beginning clozapine or risperidone, baseline antipsychotic medication was gradually decreased. Treatment was continued for up to 29 weeks with maximum dosage of 800 mg/day of clozapine or 16 mg/day of risperidone after 5 weeks.

Statistical Analyses

Changes in spatial working memory performance were examined by computing mixed effects regression models (21) with drug (clozapine or risperidone) as the independent variable, baseline spatial working memory performance as a covariate, and

### Table 1. Demographic and Clinical Characteristics of Patients With Moderately Refractory Schizophrenia Randomly Assigned to Treatment With Risperidone or Clozapine, by Treatment Group and Trial Completion Status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Risperidone (N=50)</th>
<th>Clozapine (N=47)</th>
<th>Discontinued Trial (N=62)</th>
<th>Completed Trial (N=35)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>41 (82%)</td>
<td>36 (76%)</td>
<td>49 (79%)</td>
<td>28 (80%)</td>
</tr>
<tr>
<td>Female</td>
<td>9 (18%)</td>
<td>11 (24%)</td>
<td>13 (21%)</td>
<td>7 (20%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>30 (60%)</td>
<td>26 (55%)</td>
<td>36 (58%)</td>
<td>20 (57%)</td>
</tr>
<tr>
<td>Other</td>
<td>20 (40%)</td>
<td>21 (45%)</td>
<td>26 (42%)</td>
<td>15 (43%)</td>
</tr>
<tr>
<td>Education (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;12</td>
<td>18 (36%)</td>
<td>19 (41%)</td>
<td>22 (36%)</td>
<td>15 (42%)</td>
</tr>
<tr>
<td>12</td>
<td>17 (34%)</td>
<td>18 (38%)</td>
<td>25 (40%)</td>
<td>10 (29%)</td>
</tr>
<tr>
<td>&lt;12</td>
<td>15 (30%)</td>
<td>10 (21%)</td>
<td>15 (24%)</td>
<td>10 (29%)</td>
</tr>
<tr>
<td>Drug</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risperidone</td>
<td>32 (51%)</td>
<td>18 (51%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clozapine</td>
<td>30 (48%)</td>
<td>18 (48%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>41.9</td>
<td>9.2</td>
<td>21–63</td>
<td>41.9</td>
<td>8.1</td>
<td>22–58</td>
<td>41.0</td>
<td>8.2</td>
<td>43.3</td>
</tr>
<tr>
<td>Spatial working memory performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-second delay</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>23.2</td>
<td>6.2</td>
<td>4–32</td>
<td>23.7</td>
<td>5.5</td>
<td>7–31</td>
<td>23.8</td>
<td>5.4</td>
<td>23.0</td>
</tr>
<tr>
<td>29 weeks</td>
<td>24.2</td>
<td>6.1</td>
<td>12–32</td>
<td>20.9</td>
<td>7.5</td>
<td>5–32</td>
<td>20.0</td>
<td>6.8</td>
<td>20.1</td>
</tr>
<tr>
<td>15-second delay</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>20.5</td>
<td>7.7</td>
<td>0–32</td>
<td>19.9</td>
<td>6.5</td>
<td>4–32</td>
<td>20.0</td>
<td>6.8</td>
<td>20.1</td>
</tr>
<tr>
<td>29 weeks</td>
<td>21.7</td>
<td>6.2</td>
<td>9–29</td>
<td>17.9</td>
<td>6.7</td>
<td>5–32</td>
<td>12.1</td>
<td>6.9</td>
<td>10.0</td>
</tr>
<tr>
<td>Baseline psychotic symptoms</td>
<td>14.9</td>
<td>4.0</td>
<td>15.9</td>
<td>4.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline negative symptoms</td>
<td>8.3</td>
<td>2.9</td>
<td>8.6</td>
<td>3.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
spatial working memory performance at 17 weeks and 29 weeks as the repeated dependent variables. The main effect for drug in these analyses is a test of whether the two medication groups differed in spatial working memory performance after baseline performance was controlled. Analyses were conducted by using SAS PROC MIXED (22). Parallel analyses were performed for the 5- and 15-second delays of the spatial working memory test. In order to evaluate whether associations between symptoms and cognitive functioning accounted for the observed drug effects, BPRS psychotic symptom and anergia subscale scores at the two assessment points were included as changing covariates in two subsequent mixed effects models with the same independent and dependent variables.

Results

Subjects receiving clozapine and risperidone did not differ in terms of any background characteristic or on measures of baseline psychopathology and spatial working memory, as determined by t tests and chi-square analyses (Table 1). A total of 64 subjects (65%) completed the 17-week assessment, 46 subjects (47%) completed the 29-week assessment, and 35 subjects (36%) completed all three assessments. Comparisons between study completers and noncompleters on background characteristics, symptoms, and cognitive performance indicated no significant differences (Table 1). Seven subjects with baseline values only were excluded from the analysis. The mixed effects regression models indicated a significant main effect for drug for spatial working memory at the 5-second delay (F=6.60, df=1, 61, p<0.02) and the 15-second delay (F=13.52, df=1, 61, p=0.0005). The addition of the covariates of psychotic symptoms and anergia to the models had negligible effects on the main effect for drug (5-second delay: F=6.76, df=1, 58, p<0.02; 15-second delay: F=13.76, df=1, 58, p=0.0005). None of the drug-by-time interactions was statistically significant (p>0.10). As can be seen in Figure 1, the main effects for drug reflect a pattern in which the spatial working memory performance of patients receiving risperidone improved from baseline at both the 5- and 15-second delay conditions, whereas the performance of patients receiving clozapine worsened.

Discussion

In this 29-week, double-blind evaluation of the effects of clozapine versus risperidone on spatial working memory, the performance of patients receiving risperidone improved and that of patients receiving clozapine worsened. The differential effects of clozapine and risperidone on spatial working memory performance were evident at 17 weeks of treatment and consistent through 29 weeks. The effects of clozapine and risperidone were similar for both the 5- and the 15-second delay conditions, and were actually stronger when symptoms were statistically controlled. The differential effect of risperidone and clozapine on spatial working memory performance may be due to differences in the receptor affinities of the two drugs. Specifically, clozapine has substantial anticholinergic properties that risperidone lacks (17). That the anticholinergic properties of clozapine account for the worse performance on the spatial working memory task is supported by substantial evidence that anticholinergic drugs impair this function (23) and by evidence that clozapine impaired performance on a visual memory task similar to the spatial working memory task (11). The improvement in spatial working memory performance associated with risperidone treatment may be due to its ability to enhance dopaminergic turnover in the frontal cortex in the absence of potent anticholinergic effects, consistent with the effects of experimental dopaminergic agonists on spatial working memory performance in nonhuman primates, as stated in the introduction.

The findings in the current study are consistent with those from a previous double-blind comparison of haloperidol and risperidone that used the same spatial working memory paradigm (8). In that study, haloperidol worsened performance compared with risperidone, an effect associated with benztrpine use in patients receiving haloperidol. Risperidone treatment was associated with improved performance relative to haloperidol and to baseline performance, similar to results reported here. Taken together, these two studies support the conclusion that...
risperidone improves spatial working memory, perhaps by potentiating prefrontal dopaminergic activity.

There was significant attrition in the current study, with 97 subjects participating in the baseline assessment, 62 (64%) at the 17-week assessment, and 35 (36%) at the 29-week assessment. However, there were no differences between patients who dropped out and study completers in terms of any background characteristics, baseline symptoms, or spatial working memory performance. Furthermore, there were no differences in the dropout rate between the clozapine and risperidone groups, suggesting that two drugs were comparably tolerable.

Treatment-related changes in spatial working memory performance were unlikely to be related to differential drug impact on stimuli perception or motor response because practice trials with no delay between stimuli and response were conducted before testing to ensure that subjects were able to 1) perceive test stimuli, which were relatively large (1-inch diameter circles), and 2) carry out the response (pointing to a target stimulus) with 90% accuracy. Risperidone and clozapine may have differential impact on aspects of attention necessary for task performance, a topic of interest for future studies.

In sum, clozapine and risperidone demonstrated differential effects on spatial working memory, with clozapine impairing performance and risperidone improving it. These effects appear to be due mainly to the antimuscarinic blockade produced by clozapine and the enhanced prefrontal dopaminergic transmission associated with risperidone.

References


Familial Aggregation of Suicidal Behavior: A Family Study of Male Suicide Completers From the General Population

Caroline D. Kim, B.Sc.
Monique Seguin, Ph.D.
Nadia Therrien, B.A.
Geneviève Riopel, B.A.
Nadia Chawky, M.Sc.
Alain D. Lesage, M.D., M.Phil.
G. Turecki, M.D., Ph.D.

Objective: This study compared suicidality in families of adult male suicide completers and community comparison subjects.

Method: Two hundred forty-seven relatives of 25 male suicide completers and 171 relatives of 25 matched comparison subjects were assessed for recurrent risk of suicidal and related behaviors. Analyses were performed on a subgroup of relatives of suicide completers with cluster B personality disorders.

Results: Relatives of suicide completers were over 10 times more likely than relatives of comparison subjects to attempt or complete suicide after the authors controlled for psychopathology. Relatives of suicide completers were not more likely to exhibit suicidal ideation but had more severe suicidal ideation than relatives of comparison subjects. These findings were stronger for the suicide completers diagnosed with cluster B personality disorders.

Conclusions: Suicide has a familial component independent of psychopathology that may be mediated by a combination of factors, including more severe suicidal ideation and aggressive behavior.

Studies have indicated a potential role for a familial factor in the etiology of suicidality (1–4). However, psychopathology, arguably the strongest predictor of suicide, also clusters in families, and the nature of the relationship between these two factors remains unclear. Familial aggregation of mental illness and familial aggregation of suicide are not entirely causally related (5–9), and family studies in adolescents have shown that even after adjustment for psychopathology, relatives of completers and attempters have higher rates of suicidal behavior than comparison subjects (10, 11). Further studies have suggested a role for impulsive aggression in families of mood-disordered suicide attempters and nonattempters (12, 13). However, these results have yet to be confirmed in unselected suicides.

Here we report preliminary results on a family study that confirms the familial transmission of suicidal behavior independently of psychopathology in the first- and second-degree relatives of adult male suicide completers from the general population. Our results also suggest the importance of the severity of suicidal ideation and aggression in the familial transmission of suicidality.

Method

The subjects for this study were 25 male suicide completers who were randomly selected over a 3-year period from the general population of Montreal and 25 age- and gender-matched comparison subjects. The suicide completers were identified through a collaborative agreement with the Quebec Coroner’s Office, and family members appearing at the morgue were invited to participate in the study.

Community comparison subjects were recruited with the best-acquaintance method, whereby comparison subjects were randomly selected from five friends of the deceased. For cases in which no matched friend was available, the recruitment of comparison subjects was done by local advertisement. Comparison subjects were excluded if they had a history of a major axis I disorder. This study was approved by our local institutional review board, and written informed consent was obtained from participating family members after the study had been fully explained.

Consensus psychiatric diagnoses for probands were made by means of the psychological autopsy method with DSM-IV criteria (14). Structured interviews were carried out by trained clinicians with informants best acquainted with the subject by using modified Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS) interviews, including questions from the Interview Schedule for Children to assess personality disorders (15). To assess how K-SADS interviews would compare to Structured Clinical Interviews for DSM-IV in this adult population, 15 subjects were randomly assessed with both instruments. Agreement for major psychiatric diagnoses was 100%. Procedural details and reasons for using the K-SADS in this particular population are discussed elsewhere (16, 17). Interrater reliability for diagnoses was estimated in a subgroup of subjects, and kappa coefficients for key psychopathology were above 0.80 (16).

Psychopathology of relatives was assessed through administration of the Family Interview for Genetic Studies (18). Suicidal behavior and suicidal ideation (the latter was obtained with an adapted version of the Scale of Suicide Ideation [19] providing a measure of current ideation) were assessed directly with relatives. Relatives were all first- and second-degree relatives of suicide completers and comparison subjects who either were directly assessed for suicidality (suicides, 55.9%; comparison subjects: 53.2%) and/or were described by an informant during the course of the Family Interview for Genetic Studies. Variables were defined according to information obtained through the Family Interview for Genetic Studies and were assessed by the same rater.
TABLE 1. Psychiatric and Psychological Traits of Relatives of Suicide Completers and Relatives of Community Comparison Subjects

<table>
<thead>
<tr>
<th>Trait</th>
<th>Relatives of Suicide Completers</th>
<th>Relatives of Comparison Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First-Degree (N=113)</td>
<td>Second-Degree (N=134)</td>
</tr>
<tr>
<td>Suicidal behavior</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>Illegal drug use</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Alcohol problems</td>
<td>20</td>
<td>42</td>
</tr>
<tr>
<td>Aggression</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Anxiety</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

*p≤0.05.  **p≤0.01.  ***p≤0.001.

History of depressed mood, illegal drug use, and alcohol problems were all significantly more likely to occur in the relatives of suicide completers than in the relatives of the comparison subjects. There was a significant difference in the presence of aggressive behavior in the first-degree relatives (odds ratio=3.97, 95% confidence interval [CI]=1.23–12.80; p<0.02). When the first- and second-degree relatives were considered, the effect was less intense, with a nonsignificant tendency for aggressive behavior in the relatives of suicide completers (6.1%) in relation to the comparison subjects (4.1%). Other Family Interview for Genetic Studies personality and psychiatric traits are shown in Table 1.

A total of 227 relatives were directly assessed for suicidal behavior and ideation (relatives of the suicide completers=132, relatives of the comparison subjects=95). The relatives of the suicide completers and the relatives of the comparison subjects were not significantly different for past or present suicidal ideation (odds ratio=2.08, 95% CI=0.98–4.42). However, the relatives of the suicide completers had significantly higher suicidal ideation scores than the relatives of the comparison subjects (t=-2.69, df=225, p=0.008).

Previous work in our group has suggested that suicide completers can be divided into subgroups according to psychopathology, most notably by the conditions characterized by impulsive and aggressive traits (17). Thus, a subgroup of suicide completers with cluster B disorders was compared to the comparison group. Thirteen suicide completers were diagnosed with cluster B personality disorders.

Aggression in the first-degree relatives of cluster B completers was significantly higher in relation to the relatives of comparison subjects (odds ratio=5.43, 95% CI=1.62–18.19; p<0.006). When all relatives were considered, the tendency remained, although it became nonsignificant (7.6% of the relatives of cluster B completers versus 4.1% of the relatives of the comparison subjects). The adjusted odds ratio for the effect of family history on suicidal behavior was 14.84 (95% CI=1.82–120.81; p<0.05). Depressed mood had an adjusted odds ratio of 4.90 (95% CI=1.55–15.46; p=0.01). Neither alcohol problems (odds ratio=1.15, 95% CI=0.32–4.10) nor illegal drug use (odds ratio=4.01, 95% CI=0.85–18.88) remained significant in the logistic regression model.
Discussion

We found that a family history of suicide remains a significant predictor of suicidal behavior even when psychopathology is taken into account and that aggression is significantly higher among the first-degree relatives of suicide completers than among the relatives of comparison subjects. When we looked at a subgroup of suicide completers with cluster B personality disorders, we found that the effect of family history in predicting suicidal behavior was strengthened both in intensity and in significance, as was aggression in first-degree relatives. The latter finding is interesting, as previous studies have found that higher proband aggression is associated with higher familial suicidal behavior in adolescent completers (10).

We also found that the occurrence of suicidal ideation is not significantly different between the relatives of completers versus the relatives of comparison subjects. However, we found that the relatives of completers have higher levels of ideation than do the relatives of comparison subjects, suggesting that it is the severity rather than the presence of suicidal ideation that may be part of the liability toward suicide transmitted in families. This is consistent with models proposed to understand the predisposition to suicide (20).

Of interest, the relatives of suicide completers—in both the large group and the subgroup—were more likely to be deceased than the relatives of comparison subjects, even though the relatives were not different in age, gender, or the presence of physical or prolonged illness. It is possible that the relatives of suicide completers are engaged in activities or behaviors that put them at a higher risk for death. However, this should be confirmed and further investigated.

Nonsignificantly higher rates of criminal behavior and sexual abuse were found in relatives of both groups of suicide completers. This is consistent with previous studies linking suicidality with criminal behavior (21) or a history of childhood sexual abuse (12). The lack of significance of these tendencies may be due to the power limitations of our study. Further work would benefit from more thorough assessment of cluster B personality disorders and aggressive traits.

In conclusion, we have confirmed that suicidal behavior has a strong familial component and that this family effect is particularly high in the relatives of suicide completers with cluster B personality disorders. Our results further suggest the possibility that aggressive behavior together with severe suicidal ideation may be factors in the familial transmission of suicidal behavior.

References

Quetiapine Discontinuation Syndrome

To the Editor: We describe a case of incapacitating quetiapine withdrawal, its treatment, and possible causal mechanisms.

Ms. A, a 36-year-old woman with rapid-cycling bipolar II disorder and premenstrual mood exacerbation, was treated as an outpatient with lamotrigine, 400 mg at bedtime, and clonazepam, 0.5 mg t.i.d. Quetiapine, 100 mg at bedtime, was added for residual symptoms. Although the treatment was efficacious, Ms. A gained 20 lb in 6 months and asked to stop the medication. She was advised to decrease her quetiapine to 50 mg at bedtime. After 1 day, she reported nausea, dizziness, headache, and anxiety severe enough to preclude her normal daily activities. She was instructed to take quetiapine, 75 mg the next night, but her intolerable symptoms continued. They resolved when she returned to a dose of 100 mg at bedtime. A slower quetiapine taper of 12.5 mg every 5 days with the antiemetic ondansetron also failed. On a third attempt, prochlorperazine successfully reduced her discontinuation symptoms, although moderate nausea persisted for 2 days after the taper was completed. No other medications were changed, so quetiapine withdrawal was the most likely explanation for Ms. A’s symptoms.

To our knowledge, one previous case of quetiapine withdrawal has been reported (1). An inpatient with schizophrenia experienced nausea, emesis, lightheadedness, diaphoresis, orthostasis, tachycardia, and nervousness after abruptly stopping quetiapine, 300 mg/day. These symptoms resolved when quetiapine was restarted and cross-tapered to risperidone.

Three neurotransmitters may play a role in this discontinuation syndrome. Quetiapine is a dopamine D2, serotonin 5-HT1A, and histamine H1 receptor antagonist. Dopamine, serotonin, and histamine receptors are present in the chemoreceptor trigger zone, a medullary site that causes nausea and emesis when stimulated. Dopamine and serotonin also influence autonomic control in brainstem nuclei. Therefore, these neurotransmitters are present in brain regions that could cause nausea and autonomic dysregulation. Ms. A’s symptoms were similar to those reported in patients who were withdrawn from other atypical (2) and typical (3) antipsychotics, implicating dopamine. They also resembled the selective serotonin reuptake inhibitor discontinuation syndrome (4), suggesting a role for serotonin. However, ondansetron, a 5-HT3 antagonist, was ineffective in controlling nausea in this case, whereas prochlorperazine, an H1 and D2 antagonist, mitigated the severest symptoms. These clinical findings and the fact that quetiapine affects predominantly histamine at low doses suggest that Ms. A’s symptoms were due to withdrawal of H1 antagonism. However, her therapeutic improvement at only 100 mg/day raises the possibility that she was highly sensitive to all of quetiapine’s pharmacological effects.

This unusual case demonstrates that quetiapine may cause significant discontinuation symptoms in susceptible individuals. Prochlorperazine may attenuate these symptoms, allowing for successful weaning.

References

Low IQ and Gasoline Huffing: The Perpetuation Cycle

To the Editor: There is relatively little information about the long-term cognitive effects of long-term, recreational gasoline inhalation among functional subjects who are still residing in their communities (1). We present a case of severe cognitive impairment secondary to long-term gasoline inhalation with a background of low intellect.

Mr. A, a 19-year-old man, was seen for the second time with psychosis and bizarre behavior secondary to long-term gasoline inhalation over the last 2 years. To enhance the effects, he would heat the gasoline container under a hot water tap or in the oven and would add washing detergent and medicines, hoping to get an extra “high” from sniffing this concoction.

Mr. A had a developmental delay in childhood, especially in gross motor and language skills. School reports reflected that very poor performance and polysubstance abuse were prevalent from his teenage years. Cognitive testing at age 16, before recreational gasoline inhalation, demonstrated global impairments and placed him within the mildly mentally deficient category on the Wechsler Intelligence Scale for Children (WISC-III). A second cognitive assessment on the Wechsler Adult Intelligence Scale (WAIS-III) (2) was administered during this admission. The results he obtained on the verbal, performance, and full-scale IQ all fell within the extremely low category of cognitive functioning. It was clear that there had been a considerable decrease in intellectual functioning since the first assessment.

Since there was no obvious structural brain deficit, a single photon emission computed tomography (SPECT) scan was performed to ascertain whether there was functional impairment. It showed a global decrease in brain activity, especially in the frontal lobes. The anterior edge had such low activity that on a coronal section, they were markedly flattened, with a “Swiss cheese” appearance. This finding was consistent with the poor executive functioning exhibited on the WAIS-III. There was also decreased activity in the inferior orbital lobes, the temporal lobes, and the cerebellum. It is unclear to what extent the SPECT scan abnormalities were related to the developmental delay in childhood or the effects of long-term gasoline inhalation. However, with the extent of frontal lobe functional impairment, it is likely that this was significantly worsened by gasoline inhalation abuse. This might explain the poor impulse control and judgment, which worsened Mr. A’s addiction to gasoline inhalation. Gasoline inhalation continues to perpetuate frontal lobe dam-

References
age, creating a vicious cycle of increasing addiction and brain damage.

A magnetic resonance imaging scan would likely provide only equivocal information on the extent of brain damage because brain mass remains constant, despite the continued loss of brain functional capacity. The SPECT scan was useful for outlining the functional capacity of the brain and might be a useful adjunct in the future for assessing brain impairment from substance inhalation.

References
2. Wechsler D: Wechsler Adult Intelligence Scale, 3rd ed. San Antonio, Tex, Psychological Corp (Harcourt), 1997

LEONA YIP, M.B.Ch.B.
AHMED MASHHOOD, M.B.B.S.
SUNÉ NAUDÉ, Ph.D.
Wellington, New Zealand

Anticonvulsant Hypersensitivity Syndrome From Addition of Lamotrigine to Divalproex

TO THE EDITOR: Lamotrigine is an efficacious, well-tolerated treatment for bipolar disorder and seizure disorders. The initial dosing requires gradual dose escalation, especially when treatment for bipolar disorder and seizure disorders. The incidence of anticonvulsant hypersensitivity syndrome related to lamotrigine also was taking valproic acid (2). The overall rate of rashes for patients taking lamotrigine is 13% and of serious rashes, 0.1% (3). Any rash is potentially serious and should be evaluated promptly (4). Although prolonged symptoms and fatalities have been reported, early recognition and discontinuation of offending agents often result in rapid improvement, as with our patient.

References

MOHAMED RAHMAN, M.D.
NADEEM HAIDER, M.D.
Fargo, N.D.

Bipolar Disorder and Niemann-Pick Disease Type C

TO THE EDITOR: A number of neurodegenerative disorders that appear in adolescence or early adulthood are associated with the development of major mental illness. We present the case of a young man who was seen with a bipolar illness in the setting of the early-life diagnosis of Niemann-Pick disease type C.

Mr. A was a 25-year-old man who was hospitalized for behavioral disturbances, including disturbed sleep and appetite, sexualized behavior, and disorientation. Born prematurely, he was diagnosed with jaundice and hepatosplenomegaly. Fibroblast testing confirmed deficient cholesterol esterification and positive filipin staining, and later mutation analysis revealed homozygosity for the I1061T mutation of the NPC1 (Niemann-Pick disease type C) gene. He attended a special educational program and was able to work and participate in musical theater. At age 18, he exhibited early cognitive decline, and at 23, his behavior became disturbed, with periods of increased wakefulness, elation, poor judgment, and disinhibition that lasted 2–3 weeks, every 1–2 months, with interepisode euthymia.

At a mental status examination, Mr. A was highly elevated, euphoric, and markedly sexually disinhibited, with sexualized thought content. There were no delusions, for-
mal thought disorder, perceptual abnormalities, or ideas of harm. He was disoriented and distractible, and his judgment was poor. He showed dystonic hand posturing, although his power, reflexes, and sensations were normal. His speech was dysarthric, and he had limited upward and no downward saccades, although his lateral eye movements were preserved. He showed mild hepatosplenomegaly.

Magnetic resonance imaging of Mr. A’s brain showed mild generalized atrophy, prominent in cerebellar hemispheres and anterior temporal areas. His visual evoked potentials were normal. His full-scale IQ on WAIS-III-R testing was 64. He was diagnosed with bipolar I disorder, with rapid-cycling features. Initiation of sodium valproate, 500 mg b.i.d., resulted in significant stabilization of his elevated mood, leading to discharge. His improvement was sustained over 18 months of follow-up.

Niemann-Pick disease type C is an autosomal-recessive disorder of cholesterol metabolism that has been reported to appear as a schizophrenia-like illness up to a decade before the onset of its characteristic cognitive decline, ataxia, and vertical supranuclear ophthalmoplegia (1, 2). It is caused by a mutation to either of the NPC1 or NPC2 genes, which are involved in intracellular cycling of sterols (3), causing an accumulation of cholesterol in neurons and resulting in axonal and neuronal loss in callosal, cerebellar, and hippocampal regions (4). To our knowledge, this is the first report of a bipolar-type illness in Niemann-Pick disease type C adding to the established literature suggesting that up to 40% of patients with adult-onset illness are seen with psychosis.

References

DANNY SULLIVAN, F.R.A.N.Z.C.P.
MARK WALTERFANG, F.R.A.N.Z.C.P.
DENNIS VELAKOULIS, F.R.A.N.Z.C.P.
Melbourne, Victoria, Australia

Suicide and Latitude in Argentina: Durkheim Upside-Down

To the Editor: Links between suicide, seasonal affective disorder, and latitude are known. We present new and potentially useful information linking seasonal light variation to suicide. To our knowledge, there have been no previous reports of the association between latitude and suicide rates in the extreme southern hemisphere. We found no studies in MEDLINE that included countries south of 45° south latitude. Argentina is the only country with an appreciable population below 45° south latitude. We analyzed suicide rates in the three southernmost provinces of Argentina (Chubut, Santa Cruz, and Tierra del Fuego); suicide rates were 5.9 (in women) and 21.4 (in men) per 100,000 compared to national rates of 3.1 (in women) and 15.5 (in men) (1). Native American population distribution in Argentina shows that the northernmost province, Salta, has the highest percent of Native Americans (15%), far higher than the national average of about 0.5% (Instituto Nacional de Asuntos Indigenas, Buenos Aires). This information should encourage researchers studying the effect of seasonal or diurnal light variation on suicide rates.

Since Durkheim’s work in 1897 (2), there have been many studies linking latitude and suicide. The latitude-suicide connection is also shown by data from 12 U.S. Indian Health Service administrative areas. Alaska has close to the median rates for homicide and unintentional injury but has three times the median suicide rate (3).

Explanatory factors may include race, socioeconomic differences, temperature, and diurnal or seasonal light variation. Examples of ethnic differences are higher rates in Hungary than in England and in Finland than in Sweden at the same latitudes. The common factor of Finno-Ugrian ethnicity has been proposed. The suicide rates in Native Americans are higher than in the U.S. general population (3).

A possible neurophysiological explanation for the suicide-latitude link is the recent discovery of retinal melanopsin receptors directly linked to the circadian rhythm center in the suprachiasmatic nucleus (4). The extreme seasonal variation of light and dark at high latitudes may be linked to high suicide rates and seasonal affective disorder and suggests the possibility of preventive light therapy.

Further high latitude studies, both north and south, could explore potential prevention of suicide and seasonal depression.

References

ALICIA E.B. LAWRYNOWICZ, M.D., M.P.H.
Mar del Plata, Argentina
TIMOTHY D. BAKER, M.D., M.P.H.
Baltimore, Md.

Olanzapine Augmentation of Fluoxetine in Body Dysmorphic Disorder

To the Editor: This letter reports on olanzapine augmentation of fluoxetine in six patients with body dysmorphic disorder. Body dysmorphic disorder is a relatively common and severe disorder whose pharmacotherapy has been only minimally studied (1). Body dysmorphic disorder appears to often respond to serotonin reuptake inhibitors (SRIs) (1–3), but most patients do not respond or respond only partially. Investigation of SRI augmentation strategies is therefore needed. Because 35%–50% of patients with body dysmorphic disorder are delusional (1), SRI augmentation with antipsychotics is of particular interest.
An institutional review board approved this study, and the subjects provided written informed consent. Six subjects (50% women, mean age=29.3 years, SD=11.9) were first treated with fluoxetine for ≥12 weeks (mean dose=70.0 mg/day, SD=11.0). Olanzapine was then added to fluoxetine (the fluoxetine dose was unchanged) if the fluoxetine response was inadequate (i.e., the subjects still met DSM-IV body dysmorphic disorder criteria, had a body dysmorphic disorder score ≥20 on the Yale-Brown Obsessive Compulsive Scale Modified for Body Dysmorphic Disorder (4), and were no more than minimally improved on the Clinical Global Impression scale [CGI]). Two subjects received olanzapine under double-blind conditions; four were treated openly with olanzapine after failing to respond to placebo. Exclusion criteria were standard for efficacy studies. The subjects took no other psychotropic medications. Olanzapine was begun at 2.5 mg/day, with an attempt to raise the dose to 15 mg/day over 8 weeks if it was tolerated.

With olanzapine treatment, body dysmorphic disorder symptoms on the CGI were minimally improved in two patients and unchanged in four. Olanzapine was received for a mean duration of 5.3 weeks (SD=3.1); the mean endpoint dose was 4.6 mg/day (SD=3.3). Two patients experienced fatigue, and three gained weight.

These results must be considered preliminary because of the small sample size. Nonetheless, they are consistent with the only placebo-controlled SRI augmentation study in body dysmorphic disorder, to my knowledge, in which a typical neuroleptic (pimozide) was not efficacious (5). In what is to my knowledge the only report of SRI augmentation with atypical antipsychotics in body dysmorphic disorder (a chart review study, reference 3), only two of nine subjects responded, although the effect size was large. In one case report (6), olanzapine monotherapy was efficacious. These somewhat mixed results underscore the need for further studies of atypical antipsychotics as augmentation agents and monotherapy. Because clinical experience suggests that atypical antipsychotics can be very helpful for associated anxiety and agitation, this also warrants investigation.

References

KATHARINE A. PHILLIPS, M.D.
Providence, R.I.

Pilocarpine Treatment of Xerostomia Induced by Psychoactive Medications

TO THE EDITOR: Dry mouth (xerostomia) is a frequent complication of psychoactive medications with antimuscarinic and anticholinergic side effects. The lack of saliva is annoying to patients, impairs their ability to masticate and digest food, and is a potential source of dental morbidity, including increased risk for caries and oral infection. Pilocarpine is a cholinergic muscarinic agonist. It has been used to treat xerostomia induced in cancer patients by head and neck radiotherapy (1). It has recently been found to be effective in doses of 20 mg/day in a randomized, placebo-controlled dose-adjustment study in the treatment of dry mouth and dry eyes in patients with Sjögren’s syndrome (2). It has been used to treat dry mouth as a complication of opioid treatment (3). Toxicity has been infrequently reported (4). However, it is contraindicated in patients with angle-closure glaucoma.

We have empirically used pilocarpine in doses of 10–30 mg/day, divided into dosing of two or three times a day. We have used it with our acute psychiatric inpatients, ages 20–69, who complained of dry mouth after they had been started on psychoactive medication. These included atypical antipsychotic agents, particularly clozapine and olanzapine; anticholinergic agents, primarily benztropine; and antidepressants, particularly tricyclic antidepressants and mirtazapine.

Substantial relief of dry mouth was achieved in most patients. Side effects were mainly sweating and increased urination. We did not observe any adverse impact on psychiatric symptoms. The patients were generally pleased that their dry mouth symptoms responded rapidly, usually within 1 day, to pilocarpine treatment. Further investigation into the use of pilocarpine for the treatment of xerostomia induced by psychoactive medication seems warranted.

References

KIM J. MASTERS, M.D.
St. Simons Island, Ga.
Cancer Risk in Parents of Patients With Schizophrenia

To the Editor: Reassuring as it is to see that a recent record linkage study from Denmark by Susanne Oksbjerg Dalton, M.D., Ph.D., et al. (1) accurately replicated our previous finding of a reduced incidence of cancer in the parents of patients with schizophrenia in Finland compared with the general population (2), I was surprised to recognize that this difference vanished when parents whose children were free of schizophrenia were chosen as the comparison group instead. Thus, the authors made a point that our common finding with the general population as the reference group should be invalid, and they claimed a “healthy parenthood effect” as the critical source of bias. However, when testing our genetic hypothesis, they should have considered parents exposed to cancer risk throughout their lifetime. I am not sure whether their method of having follow-up for cancer in parents starts only at the birth of the first child or, alternatively, at the birth (or age at disease onset) of the first child with schizophrenia might have biased their finding of equal cancer risk (e.g., given that some studies find schizophrenia risk to depend on birth order, e.g., Kemppainen et al. [3]). Also, the “healthy parenthood effect” they introduced from a Danish study that found parents of children with cancer at no higher incidence than the general population (4) did not receive general support from several other population studies in parents of individuals with cancer at a younger age (5–7), and it seems counterintuitive at least since a significant proportion of cancer at a younger age is known to occur on a genetic basis. In fact, the only other retrievable large population study that compared cancer risk between the relatives of patients with cancer and the relatives of otherwise deceased persons from Utah (instead of the general population) still found familial cancer risk increased (8). Therefore, I doubt the general validity of the “healthy parenthood effect.” It would have been useful to see whether cancer risk in the Danish comparison group of parous individuals with no child affected by schizophrenia was indeed lower than in the general population, including parous and nonparous individuals. Unfortunately, the lack of resources does not currently permit me to replicate, in turn, a selection of a parous comparison group from the Finnish population register and analyses similar to those of the Danish study. Meanwhile, I commend the colleagues from Denmark for drawing benefit from the excellent epidemiological material available in Nordic countries, and I hope that other appropriate registers (e.g., Hemminki et al. [9]) will contribute.

References

Dirk Lichtermand, M.D.
Bonn, Germany

Dr. Dalton and Colleagues Reply

To the Editor: We introduced the healthy parenthood effect as a proposed explanation for the observed change in results according to which comparison group was used in our study of cancer risk in the parents of schizophrenic offspring. We compared the cancer rates in the parents of schizophrenic offspring (exposed group) to those of other parents, which we consider to be the most correct, and found no difference in cancer risk. When we included persons who had not had any children and used the general population rates of cancer as the comparison, a method similar to that of Dr. Lichtermand and colleagues, we also observed a reduced risk of several forms of cancer, in line with the findings of our Finnish colleagues. It is unlikely that our results were biased by the choice of start of follow-up because we followed up from the time of birth of the schizophrenic offspring in the exposed group in both analyses. We do not think that the results published by our Finnish colleagues are specific to the parents of schizophrenic patients but more generally to being parents. The healthy parenthood effect denotes selective processes that lead to the forming or initiation of a family and, second, to the maintenance of a relatively regular and healthy lifestyle when living a family life. This would probably mean that parents, compared to all adults, smoke less, drink less, exercise more, and so forth. The results that we highlight from the literature in our article as supporting the notion of a healthy parenthood effect include mainly cancer forms with a large environmental component. To the best of our knowledge, there have been no studies of the risk of cancer in parents in general and, as Dr. Lichtermand points out, most studies of cancer risk in cancer families will reflect the high-risk nature and report the higher risk of some forms of cancer with a large familial component. However, the study by Westergaard et al. (1996)—apart from reporting a high risk of testicular cancer in the fathers and brothers of testicular cancer patients—did indeed also find a reduced risk of overall cancer, a reduction mainly carried by reduced lung cancer and gastrointestinal cancer risk. The differences in cancer risk based on whether the exposed group is compared to only other parents or the total population in Denmark must somehow be connected to...
Irritability and Depression

TO THE EDITOR: Manics are irritable. Some depressives are irritable. Ergo, some depressives are bipolar. Not necessarily true. This topic was discussed in a recent article by Giovanni B. Cassano, M.D., et al. (1).

DSM-III and DSM-IV turned diagnoses into symptom checklists. This may increase the reliability of diagnosis, but it does not follow that the symptom necessarily is associated with the diagnosis. Many diagnoses are associated with irritability or distractibility, and even more are associated with impaired concentration or insomnia. These symptoms may occur in mania, but they are not nonspecific that they cannot be said to imply mania.

Much of medicine used to be like psychiatry, i.e., without definitive diagnostic tests. Imagine diagnosing a myocardial infarction without ECGs or enzymes; a constellation of symptoms, including chest pain, diaphoresis, dizziness, irregular heartbeat, etc., suggest a myocardial infarction, but none of these symptoms alone would indicate a myocardial infarction. All occur much more frequently in other conditions.

An unfortunate (and unintended) legacy of DSM-III and DSM-IV is the attribution of diagnostic significance to nonspecific symptoms that are only diagnostically meaningful when they are part of a constellation of symptoms or a syndrome. This has led to agitated, irritable depressives being called bipolar (often “mixed”) and to the overdiagnosis of bipolar disorder (analogous to the overdiagnosis of schizophrenia prior to 1970).

Undeniably, some apparent unipolar depressives will turn out to be bipolar. However, the majority of unipolar depressives will never become manic or hypomanic, even with antidepressants, and the presence of irritability, agitation, and other nonspecific symptoms associated with mania does not make these patients even a little bit bipolar.

Reference


JEFFREY A. MATTES, M.D.
Princeton, N.J.

Dr. Cassano Replies

TO THE EDITOR: In our recent article on the mood spectrum, we showed that in patients with carefully diagnosed recurrent unipolar depression, there is variability in the lifetime experience of manic-hypomanic symptoms and that increased scores on this manic-hypomanic component of our measure of mood spectrum were associated with a higher likelihood of suicidality and paranoia.

It is undeniable, as Dr. Mattes asserts, that the presence of irritability does not necessarily imply mania. Indeed, our conclusions were not based on individual symptoms but on a dimension that includes 60 items, of which only three could be construed to assess irritability. Therefore, although irritability is frequent, it is not the most prominent aspect of the manic-hypomanic component, which includes a range of mood, energy, and cognitive features.

Regarding the attribution of diagnostic significance to “nonspecific symptoms,” our intention was not to purport that unipolar patients who have a high number of manic-hypomanic features should be relabeled “bipolar.” Still, the linear relationship found between the depressive and the manic-hypomanic components in patients with both unipolar and bipolar disorder when we used a dimensional approach suggests continuity between these disorders. Moreover, we found an association between the manic-hypomanic component and suicidality and paranoia both in unipolar and bipolar patients. In our view, this finding has important clinical implications. The question of whether this dimensional spectrum approach will eventually lead to the identification of a distinct phenotype of unipolar patients presenting similarities with bipolar patients is still open. We are currently conducting a clinical trial that we hope will shed some light on this issue.

GIOVANNI B. CASSANO, M.D.
Pisa, Italy

Sertraline for Recurrent Major Depression

TO THE EDITOR: Jean-Pierre Lépine, M.D., and his colleagues (1) evaluated the efficacy of sertraline for the prophylactic treatment of recurrent depressive disorder. We read this double-blind, randomized study with great interest and wish to raise some concerns about the methodological issues.

The use of placebo arms in randomized, controlled trials remains a controversial issue. It has been criticized on ethical grounds. In this context, the Declaration of Helsinki demands that individual patients in a study “be assured of the best proven diagnostic and therapeutic method,” even in a control group (2). This statement clearly discards the use of a placebo as a control when a “proven” treatment exists.

In this trial, the way the authors tried to establish that sertraline is more effective than placebo is misleading. Even if sertraline is worse than an existing treatment, it may still be “effective” in that it is better than no treatment (placebo). In this regard, Hill (3) pointed out that the essential medical question at issue is how the new treatment compares with the old one, not whether the new treatment is better than nothing. Similarly, Cochrane (4) stated that no new treatment should be introduced into medicine unless it has been shown in randomized, controlled trials to be superior to existing treatments or equivalent to existing treatment but cheaper or safer.

As there are drugs with proven efficacy for recurrent depressive disorders, such as lithium, we are keen to know why the authors did not try to compare the efficacy of sertraline with existing drugs. It appears that the authors were keen to reflect a drug-specific effect rather than demonstrating its relative efficacy. As readers, we would like to know why the authors carried out such a long placebo phase (2 months). The
patients were left without any medication for 2 months, and during this phase, all the potential patients who were more likely to have relapsed were dropped from the study. Sixty-one subjects (16%) discontinued treating during this phase. It clearly raised doubts as to whether the authors introduced bias at this stage by using an open placebo arm for 2 months. Additionally, the authors restricted the inclusion criteria and excluded depressed patients who had anxiety. However, in day-to-day clinical practice, we have a large proportion of patients who have depression with anxiety. The results of this study may not be applied to this group of patients. In light of these issues, it will be unreasonable to conclude that sertraline is significantly effective in the preventive recurrence of depression compared to placebo.

References

ASHLEY SMITH, F.C.PSYCH.(S.A.), M.MED.(PSYCH.), M.R.C.PSYCH.
ASHOK KUMAR JAINER, M.D., M.R.C.PSYCH.
CLAIRe BENNET, M.B. Ch.B.
Coventry, U.K.

Psychiatry and General Practitioners

TO THE EDITOR: Frederick S. Sierles, M.D., et al. (1) showed that future psychiatrists score higher than other medical students on measures of psychiatric knowledge and verbal reasoning and lower on other topics. They then concluded that more attention should be paid to the general medical education of psychiatrists. We question how the authors interpreted their data. The scores on the measures of learning were standardized. Therefore, if some people do better, then some must do worse. Given this obvious situation, isn’t it better for future medical errors because of poor training are different for psychiatrists and nonpsychiatrists. About 20% of the patients seeing primary care physicians have a significant mental disorder (2). Only 23% of the patients with depression treated by primary care physicians received an antidepressant, of whom many receive an insufficient dose (3, 4). Missed psychiatric diagnoses and undertreatment of those properly diagnosed by a primary care physician are serious problems. Is the converse true for psychiatrists? This has not been shown. Most psychiatric patients, we would imagine, have been evaluated and are treated by primary care physicians for any general medical condition before they see or are referred to a psychiatrist. Psychiatrists should advise a patient who has not been seen recently by a primary care physician to do so. Psychiatric patients routinely see primary care physicians; patients seeing primary care physicians do not routinely see psychiatrists. It seems to us that the urgent need is not for future psychiatrists to learn more general medicine but for future primary care physicians to learn more psychiatry.

References

ARTHUR RIFKIN, M.D.
Glen Oaks, N.Y.
WILLIAM DAVID RIFKIN, M.D.
New Haven, Conn.

Psychotropic Medication and Stroke Outcome

TO THE EDITOR: The number of elderly people will increase dramatically in the next few years. Many of them are given prescriptions for psychotropic drugs, such as antidepressants, antipsychotics, and hypnotics, to treat behavioral disturbances related to dementia. The same elderly patients are at a high risk of cerebrovascular events, which potentially could be exaggerated by concomitant psychotropic medications. A report (1) suggested that risperidone possibly increases the risk of cerebrovascular adverse events in the elderly. There is now a concern that discontinuation of risperidone may lead to a switch to conventional antipsychotic drugs, such as haloperidol (2), which is known to retard functional recovery after cerebrovascular events (3).

We have studied the safety of risperidone in aged male rats subjected to small cortical stroke. The experimental design tried to model clinical practice in which elderly patients who have sustained a small focal stroke or who are at a high risk of stroke are given psychotropic medication. The rose bengal model was selected to produce a cortical infarct in rats because the brain pathology is well characterized, the cortical lesion produced is consistent, and the lesion has a precise location and size. Aged rats were used since they might be more vulnerable to brain insults acutely and chronically. The rats were treated daily with risperidone before and for 3 weeks after ischemic injury. To assess neural and functional outcome, we used a new ledged-beam walking test that can detect motor deficits with a high sensitivity and a challenging match-to-place water maze test, followed by histological analyses.

Risperidone transiently worsened behavioral impairments during drug exposure but did not affect the extent of tissue damage or long-term functional outcome. This is consistent
with recent work in the *Journal* by Nathan Herrmann, M.D., F.R.C.P.(C.), et al. (4). Perhaps the alarming reports on risperidone, which could lead to use of nonatypical neuroleptics, should be reconsidered. Previous epidemiological studies (1) have been characterized by a low number of patients, polypharmacy, other diseases, compliance, and/or appropriateness of use. We suggest that additional controlled experimental approaches might better address the safety of psychotropic drugs.

References

1. Wooltorton E: Risperidone (Risperdal): increased rate of cerebrovascular events in dementia trials. CMAJ 2002; 167:1269–1270
2. Fowlie D: CSM warning on atypical psychotics and stroke may be detrimental for dementia. BMJ 2004; 328:1262

CHUAN-SHENG ZHAO, M.D.
KIRSI PUURUNEN, PHARM.D.
JUHANI SIVENIUS, M.D., PH.D.
JUKKA JOLKKONEN, PH.D.
*Kuopio, Finland*
TIMOTHY SCHALLERT, PH.D.
*Austin, Tex.*

To the Editor: Dr. Herrmann et al. reported a retrospective study concerning the risk of stroke among 11,400 patients (age >66 years) who were administered atypical antipsychotics. Use of atypical antipsychotics was not associated with an elevated risk of stroke compared with use of conventional neuroleptics.

We examined the 2-year risk of death among dementia patients who were using atypical antipsychotics or conventional neuroleptics or were nonusers. Originally, 425 patients (age >69 years) in city hospitals and nursing homes in Helsinki, Finland, participated (1); 255 had dementia and 106 had delirium, according to DSM-IV criteria.

One hundred thirty-five (52.9%) of 255 patients were given antipsychotics at baseline. 40.4% took conventional neuroleptics, and 12.5% took atypical antipsychotics. After 2 years, 118 of 255 were deceased. The death rate was 47.6% among those taking conventional antipsychotics, 21.9% among those taking atypical antipsychotics, and 50.0% among nonusers. We performed a logistic regression analysis to clarify which factors had independent prognostic value in mortality. When we entered age, gender, severe stage of dementia (clinical dementia rating=2–3), delirium, high number of comorbidities, impaired physical functioning, use of neuroleptics, use of atypical antipsychotics, and use of restraints into the model, only old age (>85 years) (odds ratio=1.71, 95% confidence interval [CI]=1.00–2.95), high number of comorbid disorders (odds ratio=1.96, 95% CI=1.03–3.73), and use of restraints (odds ratio=2.45, 95% CI=1.06–5.65) predicted mortality. It is surprising that the use of atypical antipsychotics seemed to protect against death (odds ratio=0.40, 95% CI=0.17–0.96). Conventional neuroleptics did not have any effect.

There is a concern that atypical antipsychotics increase the risk of stroke among dementia patients. In the study by Dr. Herrmann et al., there was no evidence of that. Nevertheless, the study also consisted of people without dementia. Patients with dementia are often old and frail and have comorbidities and delirious episodes with acute illnesses that possibly explain the high risk of death. They also have behavioral symptoms that are frightening for the patient and caregiver. These symptoms are the most common reason for admittance to permanent institutional care (1). Thus, these patients urgently need control for their symptoms. There are several randomized trials showing that both atypical antipsychotics and cholinesterase inhibitors are efficient in controlling these symptoms (1). The use of the cholinesterase inhibitors was quite rare—only 3%—among our participants in 1999–2000. Most patients used conventional neuroleptics. To our knowledge, there are no studies concerning the risk of taking neuroleptics among dementia patients. Our study presents that possibility.

It is possible that in our group the frailest patients were administered neuroleptics and the fittest were given atypical antipsychotics even though it would be against any recommendations. Nevertheless, our logistic regression analysis took into account comorbidities, physical functioning, age, and stage of dementia. Rather than showing an elevated risk, atypical antipsychotics had a protective effect.

Reference

1. Pitkälä KH, Laurila JV, Strandberg TE, Tilvis RS: Behavioral symptoms and the administration of psychotropic drugs to aged patients with dementia in nursing homes and in acute geriatric wards. Int Psychogeriatr 2004; 16:61–74

MINNA M. RAIVIO, M.D.
JOUKO V. LAURILA, M.D., PH.D.
TIMO E. STRANDBERG, M.D., PH.D.
REJO S. TILVIS, M.D., PH.D.
KAISU H. PITKÄLÄ, M.D., PH.D.
*Lahti, Finland*

Dr. Herrmann and Colleagues Reply

To the Editor: In response to our administrative health care database study comparing the risks of cerebrovascular accidents in elderly patients treated with either atypical or typical antipsychotics, Dr. Zhao et al. describe the lack of histological damage caused by risperidone treatment in aged rats with preexisting cortical infarcts. We applaud these investigators’ “bedside to bench” approach and find the results reassuring in light of the lack of association between the use of atypical antipsychotics and cerebrovascular accidents noted in our population-based cohort study. In the second letter, Dr. Raivio et al. report that in a small group of elderly patients with dementia or delirium, the use of atypical antipsychotics was actually associated with a 60% reduction in the rate of death. Although this study did not address the frequency of cerebrovascular adverse events, it too provides converging evidence that the risk of life-threatening events directly attributable to atypical antipsychotics may be overstated.

Concerns about the association between atypical antipsychotic use and cerebrovascular adverse events in elderly dementia patients based on data from randomized, controlled trials persist (Wooltorton, 2002; reference 1), and prescribing
information for risperidone and olanzapine have been modified to reflect this possibility. Possible biological mechanisms to account for this association might include thromboembolic effects, cardiovascular effects (such as orthostatic hypotension and arrhythmias), hyperprolactinemia leading to atherosclerosis, and excessive sedation, resulting in dehydration and hemconcentration. To date, there are few data to support these mechanisms (2). Dr. Zhao et al. suggest a number of methodological problems in the previous studies that raise the possibility of an association, including small numbers of patients, the effect of comorbid medical illness (especially preexisting cerebrovascular disease), and concomitant medication use. The latter is particularly poignant given the recent withdrawal of rofecoxib from the market because of higher rates of serious cardiovascular thromboembolic effects, including stroke. It is unclear what role such medications may have played in the randomized, controlled trials of antipsychotics.

Finally, Dr. Zhao et al. raise the concern that warnings such as those of the Committee on Safety of Medicines in the United Kingdom to avoid risperidone and olanzapine use in elderly dementia patients (3) may lead to switching to typical antipsychotics, such as haloperidol. Besides their well-documented tendency to cause both acute and long-term extrapyramidal symptoms (e.g., tardive dyskinesia), drugs such as haloperidol have been shown to negatively influence motor recovery in poststroke patients (4) and may not confer any reduced risk of cerebrovascular accidents in an elderly population relative to atypical antipsychotics (as in our article).

Further studies focusing on potential mechanisms may help to clarify the association between atypical antipsychotic use and cerebrovascular adverse events and ultimately define subgroups of patients at increased risk in whom benefits of therapy might outweigh the risks.

References
1. Wooltorton E: Olanzapine (Zyprexa): increased incidence of cerebrovascular events in dementia trials. CMAJ 2004; 170:1395

TO THE EDITOR: We read with great interest the recent report by Stephen R. Marder, M.D., and coauthors on the physical health monitoring of patients with schizophrenia (1). We agree with those authors’ concerns that “the health needs of people with schizophrenia...are not adequately addressed by clinicians in specialty mental health programs or in primary care settings.” However, we were perplexed to find that tobacco smoking was not considered in the review process, and no recommendations for monitoring and intervening on tobacco smoking were made in the report. There is a real risk that this report may worsen rather than improve the health of people with schizophrenia by advising clinicians to monitor and intervene on factors that have relatively small impact on their patients’ health while ignoring the main causes (substance use generally and tobacco use in particular).

Most of the excess mortality in schizophrenia is directly attributable to cigarette smoking (2). There is also an indirect effect of tobacco use on health through the high proportion of total income that people with schizophrenia spend on tobacco (27%) (3). Mental health professionals rarely assess their patients’ tobacco use (4), despite the existence of effective treatments (5–7).

We hope that those responsible for the care of people with schizophrenia will strive to improve the physical health of their patients but will target their efforts appropriately to the main causes of ill health and premature death rather than focusing on the side effects of psychotropic medications. Among the most important things clinicians can do for the health of people with schizophrenia is the proper assessment and treatment of nicotine dependence (8).

Tobacco Use and Cataracts in Patients With Schizophrenia

TO THE EDITOR: Although implementing many of the specific suggestions of Dr. Marder et al. for expanding the health monitoring of patients with schizophrenia could prove useful, the recommendation for biannual (for patients under age 40) or annual (for patients age 40 or older) slit-lamp examinations is
Dr. Marder and Colleagues Reply

To the Editor: Drs. Foulds and Williams make a valid point regarding our article: smoking makes a substantial contribution to the morbidity and mortality of individuals with schizophrenia. Although we decided at the onset of our consensus meeting that we would not include discussion of all of the factors that should be monitored in patients with schizophrenia, we agree with the writers that psychiatrists should inquire about smoking routinely. If patients are smokers, clinicians should review the health hazards associated with smoking and the available approaches for promoting smoking cessation. Whenever feasible, clinicians should engage in attempts to motivate patients to give up smoking, and they should refer them to a specialized smoking cessation program.

The letter from Dr. Steele provides an opportunity for us to underline the importance of mental health providers ensuring that their patients with schizophrenia receive adequate eye care. The recommendation for annual eye examinations for patients over 40 is a standard of care for all individuals rather than a special requirement for individuals with schizophrenia. The participants at the consensus meeting also recommended that psychiatrists inquire about the quality of vision on an annual basis and that they inquire specifically about changes in vision, the quality of distance vision, and the presence of blurry vision. Changes in vision should lead to a referral to an optometrist or an ophthalmologist. The additional requirement for slit-lamp examinations at 6-month intervals for patients receiving quetiapine is included in its package insert. As noted in the article, the meeting participants agreed that clinicians should follow the recommendations on the quetiapine package insert until there is more definitive evidence regarding the risk of cataracts. Because patients with schizophrenia often have risk factors for lens opacities, such as diabetes, hypertension, and poor nutrition, clinicians should inquire about visual changes and ensure that guidelines for visual monitoring are followed, independent of the antipsychotic prescribed. The article also noted that the level of evidence for an association of cataracts with specific antipsychotics could not be graded. As in several other areas, we hope that more definitive information will become available, resulting in updates to package inserts, whether those updates are more restrictive or more permissive.

STEPHEN R. MARDER, M.D.
SUSAN M. ESSOCK, Ph.D.
ALEXANDER L. MILLER, M.D.
ALAN FRIEDMAN, M.D.

MT. SINAI CONFERENCE PARTICIPANTS
Los Angeles, Calif.

Alexithymia, Personality, and Psychopathology

To the Editor: We read with great interest the recent article by Hans Joergen Grabe, M.D., and colleagues (1) concerning alexithymic features as predictive factors of psychopathology in psychotic patients. The article suggested the relevance of improving emotional awareness as a major issue of therapeutic interventions in patients with mental disorders. However, some methodological concerns should moderate the interpretation of these findings. Grabe and colleagues used a linear regression technique to calculate the relative magnitude of the prediction of psychopathology of several independent variables, such as Toronto Alexithymia Scale factors, Temperament and Character Inventory dimensions, age, and gender. However, they did not integrate in their analysis a dimensional assessment of depression. Now we know from several studies that alexithymia and depression, although not totally overlapping, are strongly associated, with depression acting as a strong mediator between alexithymic features and psychopathology (2). This is especially true for the difficulties identifying feelings subscale of the Toronto Alexithymia Scale.

References

THOMAS E. STEELE, M.D.
Charleston, S.C.

LETTERS TO THE EDITOR

Am J Psychiatry 162:5, May 2005
whose correlations with depression vary between 0.42 and 0.65 (3). It is not surprising that this factor was found to be the strongest predictor of psychopathology in the study by Dr. Grabe and colleagues. The authors should have included in the linear regression analysis a dimensional measure of depression, or they should have measured with a hierarchical regression analysis the additional part of the variance of psychopathy explained by alexithymic features beyond the variance accounted for by depression, as was done by Luminet and colleagues (4). These procedures would have strengthened their results and would have allowed a comparison with previous published studies that have shown that alexithymia is truly associated with higher levels of psychopathy and acts as a negative predictor of outcome beyond the influence of depression (5).

These limitations aside, we agree with Dr. Grabe and colleagues on the relevance of the alexithymia construct in mental disorders and on the need for developing specific psychotherapeutic techniques to improve affect identification and differentiation for emotionally dysregulated subjects.

References


MARIO SPERANZA, Ph.D.
MAURICE CORCOS, Ph.D.
OLIVIER GUILBAUD, M.D.
GWENOLÉ LOAS, Ph.D.
PHILIPPE JEAMMET, Ph.D.
Paris, France

Dr. Grabe and Colleagues Reply

To the Editor: We want to reply to the important comments of Dr. Speranza et al., who suggest including a dimensional measurement of depression as an additional independent variable in regression analyses predicting the psychopathological dimensions of the SCL-90-R. They argue that the “difficulties identifying feelings” subscale of the Toronto Alexithymia Scale and the ratings of depression have especially been shown to be correlated with each other alexithymic features and psychopathology.” First, because alexithymia is conceptualized as a personality construct demonstrating a high relative stability (Luminet et al., 2001), our intention was to assess the association between alexithymia and psychopathology, with adjustment for personality dimensions (the Temperament and Character Inventory) that have been shown to explain up to 45% of the variance in alexithymia (1). Second, our data show that depression (beta=0.32) is not only associated with alexithymia but also with a broad range of psychopathologic dimensions, including anxiety (beta=0.41) and somatization (beta=0.44). Thus, anxiety or somatization could also act “as a strong mediator between alexithymia and psychopathology.” Third, one has to decide a priori whether actual psychopathology is the dependent variable or a confounder within the model.

Ignoring these three statements, we followed the statistical recommendation of Dr. Speranza et al. Unfortunately, we do not have dimensional information on depression other than the depression subscale of the SCL-90-R. Given the high internal consistency of the SCL-90-R, each subscale, e.g., depression, entered as a confounder will show major statistical effects on the dependent variable, which is also an SCL-90-R subscale in our model. Thus, when we entered depression, the three Toronto Alexithymia Scale factors, the Temperament and Character Inventory, age, and gender as independent variables and the SCL-90-R global severity index (without items assessing depression) as a dependent variable into a hierarchical linear regression model, SCL-90-R depression resulted in R²=0.785 and difficulties identifying feelings added an additional R²ch=0.030 (Fch=41.3, p<0.001) to the variance. However, it is important to consider that the correlation (Pearson) between SCL-90-R global severity index (without depression) and the depression subscale was r=0.89! Thus, the adjustment of a statistical model for, e.g., depression is more reasonable if the confounder variable is not as highly correlated with the dependent variable. This was done by Luminet et al. (2001) in predicting posttreatment scores for alexithymia with pretreatment scores of alexithymia with adjustment for scores for depression or by Grabe et al. (1) when they predicted alexithymia with Temperament and Character Inventory dimensions with adjustment for the SCL-90-R global severity index.

Reference


HANS JOERGEN GRABE, M.D.
CARSTEN SPITZER, M.D.
HARALD J. FREYBERGER, M.D.
Stralsund, Germany

Refining the Personality Disorder Diagnosis

To the Editor: The article by Jonathan Shedler, Ph.D., and Drew Westen, Ph.D. (1), provides an excellent overview of the need to refine personality disorder diagnostic criteria. We would like to add some additional data about the complex diagnostic relationships between borderline and histrionic personality disorders in clinical practice.

Andalusia is the most highly populated region in Spain (7,606,848 inhabitants) (2). The health care network in Andalusia, which serves the region’s entire population, comprises 17 psychiatric units for acutely ill patients in general hospitals.

As part of a wider study, we have been working with the Minimum Basic Data Set at Discharge From Hospitals (3), a system of hospital record-keeping that records all patient discharges taking place at Andalusian hospitals.
It has been fully operational and reliable since 1995, and we have processed data up until 2002: 81,466 admissions in psychiatric units were recorded, with 12,919 receiving a diagnosis of personality disorder at discharge. The yearly total of borderline personality disorder diagnoses as a percentage of the total number of personality disorder diagnoses was as follows: 10.03% (1995), 10.27% (1996), 13.83% (1997), 16.22% (1998), 21.41% (1999), 17.18% (2000), 17.67% (2001), and 17.56% (2002). On the other hand, histrionic personality disorder accounted for the following percentages of all personality disorders: 25.75% (1995), 29.46% (1996), 16.98% (1997), 20.52% (1998), 16.78% (1999), 17.88% (2000), 17.96% (2001), and 19.17% (2002). The percentage of borderline personality disorder increased significantly over the first part of the period under study, whereas the percentage of histrionic personality disorder decreased over the same time period until the final years of the study when both percentages stabilized (the graphic representation of these percentages is nearly symmetrical around the x-axis).

The sum of borderline personality disorder diagnoses plus histrionic personality disorder as a percentage of the total number of personality disorder diagnoses changed little over time: 35.78% (1995), 39.72% (1996), 30.80% (1997), 36.74% (1998), 38.18% (1999), 35.06% (2000), 35.06% (2001), and 36.72% (2002). Therefore, it is not at all unreasonable to think that a situation of “diagnostic transfer” has taken place over this period of time, moving from histrionic to borderline personality disorder.

One of the keys to this finding was provided by the results of Blashfield and Intoccia’s brief report (4): the amount of literature on borderline personality disorder has experienced striking growth since the 1980s (after DSM-III), whereas the literature on histrionic personality disorder has declined since the mid-1970s. This situation has consequently influenced the practice of clinicians, including those in European countries.

New elements and conclusions can also be drawn from the excellent study by Drs. Shedler and Westen (1): On one hand, as the authors interpreted the data, clinicians may not have clearly differentiated conceptions of the two personality disorders, leading them to be confused because of their high level of comorbidity; or, on the other hand, current diagnostic classifications should perhaps not maintain their current distinction between borderline and histrionic personality disorder.

**References**


J. MIGUEL PENA-ANDREU, M.D., Ph.D.
ANDRES FONTALBA, M.D.
MONICA CARRERA, M.D.
J. LUIS MARTINEZ, M.D., Ph.D.
M. ISABEL RODRIGUEZ, M.D., Ph.D.
Malaga, Spain

**Comment on Hoarding**

To the Editor: The intriguing report by Sanjaya Saxena, M.D., et al. on hoarding (1) prompts me to write concerning what I believe is a significant defect in DSM-IV-TR regarding persons who primarily present with hoarding.

My own experience with hoarders is based on having performed close to 800 guardianship evaluations over the last 15 years. Guardianship evaluations are generally performed as home visits. The high prevalence of hoarding in this group is notable.

Dr. Saxena et al. included hoarders within obsessive-compulsive disorder (OCD). Our diagnostic manual defines compulsions as “repetitive behaviors” and gives examples that all have an active motor or cognitive component. The actions or mental acts “reduce anxiety.” Hoarding, I believe, usually arises from inaction. The only anxiety I have observed is when hoarders are threatened by forced clearing. The problem is not so much that these people collect. The problem is that they are unable to discard.

Hoarding in DSM-IV-TR appears only in the description of obsessive-compulsive personality disorder. The “diagnostic diagnosis” for obsessive-compulsive personality disorder does state that OCD should be diagnosed when hoarding is “extreme.” Thus, “extreme” hoarding is OCD, and less than “extreme” hoarding is obsessive-compulsive personality disorder. The clinician is further frustrated by being referred to the OCD section but finds the DSM-IV-TR section on OCD totally silent in respect to hoarding. Although DSM-IV-TR accommodates a dual diagnosis of OCD and obsessive-compulsive personality disorder, it clearly discourages diagnosing both conditions: “the clinical manifestations of these disorders are quite different” (p. 462).

One of the significant findings by Dr. Saxena et al. is that there is an anatomic locus for the syndrome of hoarding. This certainly should lead us to conclude that the one place hoarders meet criteria for obsessive-compulsive personality disorder. Only 15%–45% of hoarders meet criteria for obsessive-compulsive personality disorder (1, 2). No correlation was found between hoarding and...
severity and scores on an obsessive-compulsive personality disorder scale (3). Instead, hoarding behaviors correlate significantly with other OCD symptoms (3).

Dr. Amdur suggests that compulsive hoarders may not meet DSM-IV-TR criteria for OCD because they appear to suffer more from their inaction than from repetitive behaviors. However, compulsive hoarders have been found to have many repetitive acquisition and saving behaviors (3). While DSM-IV-TR does not explicitly mention compulsive hoarding as a symptom of OCD, hoarding and saving-related obsessions and compulsions are quite common in patients with OCD (4, 5) and are included in standard OCD assessments, such as the Yale-Brown Obsessive Compulsive Scale symptom checklist (6). Dr. Amdur observes that compulsive hoarders do not appear to have anxiety unless they are threatened by forced clearing, but several studies have found elevated anxiety levels in compulsive hoarders, even in comparison to nonhoarding OCD patients (7, 8). Moreover, hoarders also report a high prevalence of nonhoarding obsessions and compulsions (2, 3) and do not differ from nonhoarding OCD patients in the number of OCD symptoms reported on the Yale-Brown Obsessive Compulsive Scale checklist or overall score (7, 8).

Nevertheless, the question remains: Should compulsive hoarding be classified as a symptom subtype of OCD or as a separate disorder that is frequently comorbid with OCD? Factors that support a separate classification include significant differences between hoarders and nonhoarding OCD patients in comorbidity and family history (2), response to serotonin reuptake inhibitors (9), and the distinct functional brain abnormalities found in our recent study. Neuropsychological, brain imaging, and genetic studies comparing compulsive hoarders to nonhoarding OCD patients will be required to resolve this issue.

References


SANJAYA SAXENA, M.D.
Los Angeles, Calif.
As the general population (and by inference, each of us) anticipates an increasingly longer lifespan, the problem of dementia also looms larger and may reach epidemic proportions at its present rate of growth. Thus, we anxiously await dividends from an expanding body of work in the fields of genetics, molecular biology, and neuroimaging that may reveal the cause(s) and effective treatment(s) for this condition. In the meantime, a cohesive strategy is required to care for those suffering from serious cognitive decline.

In this context, Curran and Wattis have prevailed upon a multidisciplinary group of experts to outline a very readable and comprehensive approach to caring for those with Alzheimer’s dementia and related disorders. Their stated goal is to bridge the gap between the existing knowledge base derived from research in this area and the ability of clinicians to apply this information in a strategy that optimizes patient care. Their stated agenda is to contribute meaningfully to the process of training practitioners to identify dementia as early as possible, respect patient autonomy, and employ treatment approaches that emphasize the spiritual needs of the individual. The discussion from several perspectives results in some repetition, but, overall, the book is a sound review of our present knowledge about the dementias and their management.

The editors have divided the text loosely into three parts. The first reviews the history, definition, epidemiology, diagnosis, classification, and strategies for early identification. They note that age is the single most important risk factor. Family history, certain genetic variations (e.g., APOE, E4 homoyzgotes), and depression are also important predisposing factors that may increase the risk substantially. Conversely, there is preliminary evidence indicating that such disparate factors as regular physical activity; the use of wine, coffee, statins, nonsteroidal antiinflammatory drugs, certain vitamins (e.g., C and E); and higher educational attainment (as well as length of training) may reduce the risk of dementia.

The chapter authors underscore that dementia is a syndrome with multiple causes, such as vitamin deficiencies, infections, and endocrine dysfunction—conditions that may be correctible, or at least substantially improved. They highlight the three most common irreversible types (i.e., Alzheimer’s disease, vascular dementia, and Lewy body dementia), noting that accurate and early diagnosis can have a favorable impact on both treatment and prognosis. In this light, they review a number of simple early screening assessments, but they also note that there is no pathognomonic test for Alzheimer’s disease.

The second part uses case examples to clarify the role of different professionals (i.e., consultants, psychologists, physicians, nurses, and occupational therapists) in the assessment, diagnosis, and management of dementia. One chapter is devoted to each profession’s unique perspective, emphasizing the need to shift from a medical model orientation to a positive person-centered approach to care.

The third part examines the delivery care system for the dementias, focusing on the memory clinic, legal aspects, and the spiritual needs of the individual. The strength of this book is the inclusion of multiple perspectives, but the weakness is substantial redundancy. Despite the latter, I believe that this is an important consolidation of our present knowledge and a good primer for those involved in the care of the dementias.

PHILIP G. JANICAK, M.D.
Chicago, Ill.


The goal of this physician’s guide is to provide scientific and clinical knowledge about Alzheimer’s disease in comprehensible language. The editors have accomplished this with wonderfully succinct and short chapters. The focus of the presentations is to help practicing clinicians understand, diagnose, and treat Alzheimer’s disease, which is still underdiagnosed and undertreated.

The first section of the book covers the scientific background of Alzheimer’s disease. The first neuropathological hallmark of the disease is the presence of extracellular precipitations of beta-amyloid peptide. The second neuropathological hallmark is the presence of neurofibrillary inclusions composed of the tau protein. Lesions develop in the form of neurofibrillary tangles, first described by Aloys Alzheimer (1907), and threads. In addition to these hallmarks, Alzheimer’s disease brains also appear to exhibit evidence of reactive-oxygen-mediated injury (oxidative stress).

The epidemiology of Alzheimer’s disease indicates that 4%–10% of the population over 65 has the disease, and the percentage doubles every 5 years after age 65. In the United States, $174,000 is spent on each Alzheimer’s disease patient; with steadily increasing life expectancy, the number of patients is expected to rise from 9 million to 45 million by 2030. European studies indicate that women are at greater risk than men for developing Alzheimer’s disease after age 85; U.S. studies do not confirm this difference. There is no gender difference for rates and risks for vascular dementia. Interestingly, Native Americans appear to have a significantly lower rate of Alzheimer’s disease, although their aggregate rate of all dementias is similar to that of Caucasians. The economic burden of Alzheimer’s disease varies among countries; cost of illness studies indicate annual per patient costs of $6,500 (England), $24,400 (Sweden), $59,700 (Italy), and $53,300 (United States). Costs increase fourfold from the mild stage of Alzheimer’s disease to the severe stage.

The clinical assessment of Alzheimer’s disease is made by determining if there is an impairment of recent memory and at least one or more other cognitive disturbance: aphasia, apraxia, agnosia, and a disturbance in executive function.

Folstein’s Mini-Mental State Examination (1) is a reliable and sophisticated test for assessing Alzheimer’s disease. A score of 24 out of 30 is usually indicative of Alzheimer’s disease, and a score of 20 or below will correlate to substantial impairment in activities of daily living.

Neuroimaging has the potential to go beyond its traditional role of simply ruling out mass lesions and stroke to helping diagnose specific dementing diseases. Research shows that magnetic resonance imaging (MRI) can identify early structural changes caused by Alzheimer’s disease. Furthermore, molecular neuroimaging techniques (single photon emission computed tomography [SPECT] and positron emission tomography [PET]) reveal characteristic focal abnormalities in neurodegenerative diseases that are unrecognized by other methods.

Vascular cognitive impairment (previously termed senility, multi-infarct dementia, and vascular dementia) is the second most common single cause of dementia after Alzheimer’s disease. Incidence rates are about 5% in the elderly. Clinical indicators are abrupt onset, stepwise deterioration, fluctuating course, prolonged plateaus, early gait, seizure, urinary disturbance, and a history of stroke. Cognitive deficits are patchy rather than diffuse. Vascular dementia patients have a higher mortality rate than Alzheimer’s disease patients and have a median survival of 3.31 years, as opposed to Alzheimer’s disease patients, who live an average of about 5 years and up to 20 years in rare cases. Vascular patients also typically have more medical diagnoses than Alzheimer’s disease patients.

The clinical laboratory workup for dementia varies with how well the physician knows the patient. A titer for Lyme disease may be appropriate in the Northeast. HIV testing may be needed (yes, for the elderly, too). For a patient not well-known, laboratory tests including SMAC 25 should be done. Vitamin B deficiency levels should be determined because B12 deficiency may produce significant neuropsychiatric symptoms. Elevated thyroid-stimulating hormone levels carry a greater risk for dementia.

Imaging techniques such as computerized tomography or MRI are now standard practice. Structural imaging will reveal brain atrophy, besides hematomas and tumors. Serial MRI may detect decreases in hippocampal volume. Measuring glucose metabolism with [18F]fluorodeoxyglucose PET is now recognized as a useful biological marker of dementia and for distinguishing the type of dementing neurodegenerative disease.

There are now drugs available to treat Alzheimer’s disease that do not stop the degenerative process but appear to delay it for up to 24 months. The most effective of these are the acetylcholinesterase inhibitors: donepezil, rivastigmine tartrate, and galantamine hydrobromide. Each has been shown to delay cognitive decline and improve global functioning. Galantamine hydrobromide, the most recent of these drugs, appears to be beneficial with both Alzheimer’s disease and cerebrovascular disease. The newest drug, an N-methyl-D-aspartic acid (NMDA) receptor antagonist, memantine, has demonstrated improvement in individual symptoms and clinical evidence for neuroprotection. It is believed that the NMDA drug protects the brain from glutamate-mediated neurotoxicity.

There are many interesting topics and brief chapters that contribute greatly to this book by providing much detail. Included are chapters on Lewy body disease, the transitional stage to Alzheimer’s disease, and distinguishing Alzheimer’s disease from normal aging. This is an excellent reference for the geriatric practitioner.

Reference

MICHAEL R. BIEBER, PH.D., M.S., PSY, PHARM, Dallas, Tex.


In spite of the hopeful reports of better clinical stabilization and treatment outcomes with the second-generation antipsychotics, schizophrenia has remained the chronic disabling disorder it was in the days of Emil Kraepelin and Kurt Schneider. Caring for and studying chronic schizophrenia have recently gained some respectability with the new emphasis on the possibility of treating cognitive impairment as the core of the disorder, in the hope that we finally will be able to improve the patient’s functional outcome. The new generation of therapeutic studies based on a better understanding of the pathophysiology of schizophrenia has raised the hope of finally making a difference in patients’ lives, beyond treating positive and secondary negative symptoms. In the meantime, our older patients are becoming less visible. They are hidden away in nursing homes, jails, and the street. In spite of the present-day therapeutic opportunities, many patients are not optimally treated and suffer from health problems that are not adequately recognized and treated.

Our elderly patients with schizophrenia face the same conundrum as the younger ones. Psychiatrists and general practitioners often fall short in providing optimal treatment. As long as we do not make inroads against societal bias and make the disorder more “sexy” by developing better therapeutic agents, we will only be tweaking the edges of the disorder to improve the lot of patients with chronic schizophrenia. The specific impairments of schizophrenia require specialized attention that is not available to our aging patients most of the time. Our social service and mental health support systems are complicated, are frequently inadequate, and seem to require a cognitive performance far beyond what most of our patients are capable of.

Dr. Cohen has brought together an impressive group of experts and clinicians, who deal with all the aspects of schizophrenia in later life. In addition to providing practical information for the service provider, the book also reviews relevant literature about outcome and covers subjects such as schizophrenia in old age, comparisons with the younger population, medical problems of aging, epidemiology, aging, medications and their side effects, gender differences, and many other topics. The wide variety of relevant topics is both a strength and a weakness. It is not clear to me who the intended audience is. This comprehensive approach makes the book very useful and at the same time less accessible. As is unavoidable with this kind of book, it is superficial in some aspects and goes into considerable depths in others. It is a primer in cer-
tains certain ways but also loses some of its focus because different clinicians and researchers have different information needs. The book could have benefited from a chapter dealing with practical suggestions for the overextended mental health workers who have this population under their care, as well as from a table listing specific needs cross-referenced with the specific chapters. I would like to see a next edition extended into a handbook about this important topic. It should be required reading for any psychiatrist or mental health worker dealing with older patients who have chronic schizophrenia. Medical students and family members of patients may find useful information to get an idea about the complexity of the disorder.

DANIEL P. VAN KAMMEN, M.D., Ph.D.
Neshanic Station, N.J.

POSTTRAUMATIC STRESS DISORDER


This book is designed to both survey the modalities used in short-term approaches for the traumatized and provide the clinician with enough of the rudiments of five generic and seven trauma-focused approaches to begin to put them into practice. Since the majority of the literature on the treatment of trauma-related mental disorders emphasizes that the treatment of the traumatized should be gently paced, be respectful of the patient’s vulnerability, and take as much time as necessary to resolve impact of the traumatic events on the patient, this collection of short-term approaches is a unique contribution to the literature.

Many approaches to the treatment of trauma have been developed over the last two decades. Unfortunately, most remain relatively unfamiliar to the majority of psychiatric practitioners. This is an unfortunate state of affairs, because the events of September 11 have taught us that we may have to contend with large numbers of traumatized individuals at short notice. Medications are at best partial solutions for the vast majority of the traumatized, and “psychotherapy as usual” may prove less than powerful in resolving the remaining sequelae of traumatization.

Brief Treatments for the Traumatized is divided into three sections. The first addresses theoretical issues and is strong. The contributing authors generally embrace a cognitive behavior orientation. The second section discusses what are called generic therapies: cognitive behavioral treatment, narrative therapy, thought-field therapy, sensorimotor processing, and eye movement desensitization-reprocessing. The third explores trauma-focused treatments: multisensory trauma processing, neurolinguistic programming, emotionally focused therapy, brief multiple family group treatment, traumatic incident reduction, couples treatment for trauma, crisis debriefing, and the rewind technique. Some authorities would argue whether some of the treatments are appropriately classified as generic or trauma focused, but this is a minor concern. As in any multiauthor book, the chapters are not equally successful. The majority are competent and thoughtful and address their subjects well. However, some fail to achieve clarity, and some are overly ambitious—in trying to achieve too much they fall short of their objectives and leave the reader befuddled. This text does not offer an adequate discussion of the indications and contraindications for brief treatment in general or for particular brief treatments with particular patients. This is a major omission and compromises the reader’s capacity to contextualize the information the book conveys.

Although Brief Treatments for the Traumatized is neither a perfect nor a definitive text, I can recommend it strongly as an introduction to the spectrum of trauma treatments currently being developed and refined. I would strongly dispute the series editor’s remarks that the book’s descriptions of the different treatments give the reader enough knowledge to put these modalities into practice, but I believe that most of the chapters provide reasonably useful portraits of how each modality might be applied to the treatment of the traumatized. The vast majority of the approaches described as short-term can also be used in the context of the long-term treatment of patients with multiple or chronic traumatizations or can be imbricated within an ongoing psychotherapy to address specific trauma issues. One hopes that the reader will be motivated to learn one or more of these modalities in depth in order to help the traumatized more effectively.

RICHARD P. KLUFT, M.D.
Bala Cynwyd, Pa.


I wrote this book review on September 11, 2004, 3 years after the attacks on New York City and the Pentagon. In that 3-year span, I treated three patients who were within five blocks of the World Trade Center buildings at the time of the attacks, as well as survivors of Myanmar atrocities, rapes, murder attempts, domestic violence, physical violence, and kidnappings. In reaction to their problems, and the many problems facing the United States and the world, I found amazing comfort in the Iowa Hawkeyes beating the Iowa State Cyclones. Like many of my compatriots, I find myself focusing on physical and mental pursuits that are healing rather than obsessing on reopening healing scars of grief. This is the core thesis of this very skillfully and insightfully written “handbook” on PTSD overseen by Raul Silva, a veteran of September 11.

With 28 contributors from New York City and six from Lebanon, this book’s 15 chapters cover every conceivable nook and cranny of PTSD, a much ignored psychiatric condition defined in 1980 by DSM-III. Topics covered include 1) epidemiology, 2) resiliency and vulnerability factors, 3) risk factors, 4) legal aspects, 5) neurobiology, 6) etiology and pathogenesis, 7) clinical findings, 8) gender differences, 9) intergenerational links between mothers and children with PTSD spectrum illness, 10) assessment, 11) differential diagnosis, 12) childhood versus adult PTSD, 13) treatment of children exposed to trauma, 14) clinical case examples, and 15) PTSD in children and adolescents following war.
In a world where a hundred dollars does not buy a decently well-written textbook anymore and a million dollars buys a house in Los Angeles that Midwesterners would not pay $80,000 for, this compact little handbook is worth more than its weight in gold.

This is from a reviewer who has been through World War II, who has witnessed murders, rapes, and other acts of violence in a country in postwar transition, and who has been treating hundreds of victims of such atrocities since 1967.

Silva shows an uncanny sense of reverence and irreverence for traditional views of PTSD and, in the process, gives coherent meaning to the often conflicting and muddled views of this disorder, which all mental health professionals deal with. The aggregate impact of this handbook is in shedding light not only on what makes humans break down but also on what makes humans bounce back. Silva and his group, ironically, have made more sense of psychiatry in one volume than all the fragmented, topic-focused books I have read over the past 30-odd years.

By repeatedly focusing on 1) reexperiencing and the need to be retraumatized, 2) avoidance and numbing, and 3) hyperarousal, the contributors give new meaning to the cyber-space term “URL” (uniform resource locator) by allowing clinicians and patients alike to access language and information sources and solutions that actually interface and speak the same language.

For years, I have been struggling with the phenomena of “assortative mating” and “repetition compulsion neurosis.” Kowalik’s chapter on neurobiology and Linares and Cloître’s chapter on intergenerational links between mothers and children with PTSD spectrum illness made sense of all of these.

More importantly, the effectiveness of an eclectic approach to psychiatric problems versus devotion to a “one-size-fits-all” managed care mentality wins out. Once upon a time, we psychiatrists were exposed to a semiotic approach to all things human. This meant never neglecting the overarching importance of symbols in human nature. “Perception is reality” was once a mantra for most of us. This handbook takes us back to that mantra once again. Outcome studies and studies of the phenomenon called resiliency all point to the importance of engineering and shaping such symbolic reframing of experiences as the one key factor to healing—and perduring. Rising above the incomprehensible and overwhelming to emerge triumphant and stronger provides validation of Frederick Nietzsche’s “That which does not kill me, makes me stronger.”

The last chapter by Fayyad et al. is a masterful summation of the whole body of PTSD research: “It is not as bad as it sounds; and it can be as bad as we make it sound.” I give this handbook a five-star rating.

TRUCE T. ORDOÑA, M.D.
Davenport, Iowa


Some of the Marines who survived the battle of Iwo Jima never got the smell of death out of their memory. Few of these men, however, would ever consider themselves to be victims or in need of any specialized postcombat debriefing beyond what happened on the troop ships that steamed away from that island. At the outset of this book, the author notes how much has changed in our culture: “Today we expect that survivors of major terrorist attacks will be offered counseling.” The presence of mental health officers among Coalition troops deployed in Iraq underscores the change in our cultural perception about what constitutes good posttrauma professional practice.

This volume is well organized, is clearly written, and uses the current research about trauma’s impact on memory. Beginning with an overview of the clinical and cultural aspects of the disorder, Brewin moves to detailed discussions of trauma’s impact on identity, the puzzling ways in which trauma is remembered, and the debates around the false memory syndrome. He outlines the dual task for both survivor and therapist: addressing the posttrauma intrusive memories and reformulating the posttrauma identity. In a chapter titled “A Crisis of Identity,” Brewin details the variety of ways major trauma affects the self-identity of trauma survivors. Typically, survivors reinvest themselves in family, help other survivors, or demonstrate “an increased involvement with religious and spiritual issues.”

The author’s distinction between declarative and nondeclarative forms of memory provides us with a helpful way to understand how trauma continues to affect a patient’s life. Drawing on neuroscience research regarding the way memory functions, he observes how trauma’s capacity to overturn long-held assumptions is reminiscent of catastrophic interference overwhelming an established information system in ways that prevent the system from integrating the new lessons of the trauma with the older and more established patterns of declarative memory. The brain’s way of storing and making this new trauma-induced information available to the patient is by storing the information in nondeclarative memory, where it is “automatically elicited in a rather inflexible way under conditions that bear a strong similarity to the condition of the original learning.”

Finally, Brewin presents a three-step schema around which responses to survivors of large-scale trauma may be implemented. Immediate posttrauma intervention should be limited to “demonstrating safety, acknowledging the trauma, making support available to those who want it and providing information with a focus on supporting natural recovery.” The next 4–6 weeks should be spent “systematically monitoring” trauma victims “so that one can detect any failure of victims to adapt.” Finally, he urges that those victims who have “failed to adapt” receive “scientifically established interventions” that will help them recover.

Brewin asks clinicians and researchers to show “the same flexibility and resourcefulness shown by survivors” as we provide comfort and counsel to people “suddenly confronted with the unexpected, the unwanted and the unimaginable.”

DONALD D. DENTON, JR., D.MIN.
Richmond, Va.


“First, do no harm,” could be the motto of this book, a message that always bears repeating. The authors offer a critique
of critical incident stress debriefing in the aftermath of traumatic events. Although this type of group intervention has become an established practice, even mandatory in many first-responder organizations, rigorous clinical trials suggest that critical incident stress debriefing is ineffective for preventing the development of posttraumatic stress disorder. The authors are at their most persuasive and impassioned when presenting their case against critical incident stress debriefing. In the words of the editor,

It is not acceptable that early interventions for trauma be based exclusively on the understandable human need to help people who appear to be suffering or out of the motivation to promote organizational or corporate goals. (p. 6)

One problem with interventions such as critical incident stress debriefing is that most people may neither want nor need this sort of professional "help." Although extreme distress is common in the immediate aftermath of a traumatic event, most survivors will recover spontaneously, with support from the people they know and trust. Litz recommends a minimally intrusive crisis response called "psychological first aid": providing information and practical problem-solving assistance and comforting survivors without pressuring them to explore the details of the traumatic event. This sensible approach respects the privacy and resiliency of survivors.

Since only a minority of trauma survivors will develop posttraumatic stress disorder (PTSD), Litz and the chapter authors recommend targeting high-risk individuals for more intensive interventions (particularly cognitive behavior treatments). This approach has a number of limitations. First, at present we lack a simple and reliable screening method. Although numerous risk factors have been identified, no single factor is either necessary or sufficient to predict who will develop PTSD (1). Second, the authors do not consider the social implications of singling out individuals for treatment. The secondary prevention model they present is mainly derived from brief treatment of motor vehicle accident survivors, a population for whom the issues of shame, secrecy, and stigma are not particularly salient. Finally, this individualistic approach fails to address the need for repair of social relationships in the aftermath of traumatic events, despite the fact that social support is one of the most powerful predictors of recovery. (A refreshing exception is the excellent chapter by van Horn and Lieberman on treatment of infants, toddlers, and preschoolers. Here the therapeutic intervention is aimed specifically at repairing the relationship between mothers and children who have survived domestic violence.)

In general, the authors seem insufficiently aware of basic social ecology (2). Individual treatments are the only alternative proposed in place of the discredited large-group interventions. Participatory models of intervention that engage members of a traumatized community in designing their own crisis responses are simply ignored. Despite these limitations, this book will be of interest to researchers and policy makers seeking to develop an evidence-based approach to early intervention and disaster planning.

References

JUDITH L. HERMAN, M.D.
Cambridge, Mass.

SUBSTANCE/ALCOHOL ABUSE


The 12-step format, tradition of anonymity, and democratic governance of Alcoholics Anonymous (AA) are rightly credited to Bill Wilson, whom Aldous Huxley called "the greatest social architect of the twentieth century." Bill's life story has inspired no end of confessional autobiographies by alcohol and drug addicts, including one by Susan Cheever (1), the author of this book. She knows well the ravages of alcoholism, not only her own but those of her father, the illustrious author John Cheever, to whom she dedicates this book. Bill W's story is also well-known to countless alcoholics from AA's Big Book, Alcoholics Anonymous (2); the 1947 edition alone reached more than 19,000,000 people (3). Bill W has given hope to countless alcoholics who have passed the book from hand to hand as if it were a sacred document.

Bill had stayed away from alcohol until the age of 22 because of his family history of drinking problems. Having already suffered from episodes of depression, he took his first drink at 22 and experienced a glow of self-confidence and escape from dysphoria. This was followed by his descent into alcoholism and many failed attempts at detoxifications to curb his compulsive drinking. On June 10, 1935, the date affixed to the beginning of AA, Bill had been traveling on business in Akron, Ohio, and struggling to control his urge to drink. He sat down with another alcoholic, a physician, so that both could stave off their craving for alcohol, which they succeeded in doing. He later drew on his experience with the Oxford movement, which was premised on the Swedenborgian conception of steps leading to salvation, and expanded on their six-step ritual of spiritual and mutual support into the AA creed.

Cheever captures well Bill's charisma, his talents as a raconteur, his perceptiveness as a social analyst, and, to some extent, his failings (infidelity to his long-suffering wife, for example), but she understandably cannot dig deep into the social psychology underlying the AA movement. Its remarkable ability to effect changes at the mind-body interface takes place on the basis of the highly supportive welcome from the members and the interplay of cognitive changes (the acquisition of a belief system) and social cohesiveness (engagement in an intensely bonded social group). These experiences operateantly reinforce involvement, and let many members yield control of the biologically grounded pursuit of alcohol. The democratic structure of the movement and the anonymity embodied in its Twelve Traditions have served to sustain its success and hinder the emergence of self-serving leaders. Ad-
mittedly, some of the public confessionals like the one that Cheever herself published may be nibbling away at the tradition of anonymity.

Psychiatrists have a great deal to learn from AA in terms of the importance of mutual support and spiritual redemption in achieving recovery from the biologically grounded disorders that they address. This book provides a glimpse into the lives of some of the people for whom such recoveries came about. On this count, it provides a feel for what the AA experience is like and thus an opportunity to use it more effectively as an adjunct to clinical care.

References

MARC GALANTER, M.D.
New York, N.Y.

Handbook of Clinical Alcoholism Treatment, edited by Bankole A. Johnson, Pedro Ruiz, and Marc Galanter. Baltimore, Lippincott Williams & Wilkins, 2003, 316 pp., $42.00 (paper).

The editors and chapter authors of the Handbook of Clinical Alcoholism Treatment set out to provide a comprehensive, easily accessible review addressing crucial issues in the understanding of the disease of alcoholism. They have accomplished this with extraordinary attention to detail in what has morphed into a mini-textbook rather than a cursory review or handbook. For primary care providers and addiction specialists, this resource provides extensive epidemiological, neurobiological, psychological, and sociocultural knowledge relevant to the understanding and successful treatment of patients suffering from alcoholism. Alcoholism is defined as a primary, chronic disease with genetic, psychosocial, and environmental factors influencing its development and manifestations. Often progressive and fatal, it is characterized by periodic or continuous impaired control over drinking, preoccupation with the drug alcohol, use of alcohol despite adverse consequences, and distortions of thinking, most notably denial. The multifaceted etiology and natural course of alcoholism mandate a broad-based understanding and multidisciplinary approaches to assessment and treatment.

In the first two of five sections, the handbook addresses the fundamentals of alcoholism and diagnostic tools, respectively. Appropriate emphasis is placed on psychological foundations, including the important principles of social learning theory as well as the “brain disease” model of neurotransmitter-influenced reward pathways of reinforcement. Most notable is an excellent system-by-system review of the medical consequences of alcoholism with special emphasis placed on the brain and neuroimaging. Diagnostic criteria and laboratory markers are discussed and listed for easy access by the clinician. Overreliance on exclusively descriptive psychiatric diagnostic criteria and patient self-reports makes the psychiatric diagnosis of alcohol abuse or dependence challenging (1).

Research-driven advances in understanding many biological elements of alcoholism have provided the foundation for possible therapeutic medications (2). Sections three and four of this book focus primarily on specific psychotherapeutic and pharmacological modes of treatment and the variety of settings where these interventions can effectively occur (3). The authors and editors have provided a thorough and remarkably up-to-date review of these agents. However, it is not part of the dialogue to wonder why naltrexone is not effective as a treatment for alcohol dependence outside of research trials. Whether acamprosate or other treatments will be successful or will suffer a similar fate is an important question. The Handbook does not conclude what is acceptable research or provide clinical models. With others (4), one of us (M.S.G) has proposed 5-year outcomes with urine test confirmation to evaluate the role of different and often competing treatments in alcoholism.

The discussion of successful traditional psychotherapeutic approaches, including cognitive behavior therapy, motivational enhancement, and 12-step programs alone or in combination is likely more important to the clinician. The chapter authors appropriately acknowledge the importance of treatment/patient location and provide excellent sections on emergency management and primary health care settings as well as on the role of employee assistance programs and the criminal justice system. The increased awareness and understanding of the importance of identifying problem drinking combined with knowledge of treatment types, options, and locations will provide clinicians the tools to match patients to treatment properly for effective intervention and recovery.

Finally, section five provides an excellent and interesting overview of special topics, including fetal alcohol syndrome, dual diagnosis considerations, and issues specific to women, the elderly, and HIV-infected populations.

Notably missing from this fine text is a thorough description and discussion of the 12 steps of Alcoholics Anonymous (AA), their history, how they work, and what patients/clinicians can expect (5, 6). Al Anon, Alateen, and other 12-step meetings have helped loved ones understand the disease of alcoholism and obtain loved ones the treatment that they need. By leaving AA to the Big Book (7), this handbook does little to counter the argument made by addictions professionals that psychiatrists are adverse to the goals and treatment model that so many patients and family members choose. We have found that the greater our understanding of 12-step recovery, meeting types, and myths regarding AA, the more effectively we can manage resistance to this essential component of recovery from problem drinking. Psychiatrists are interested in how self-help works; AA, getting a sponsor, etc., should be a part of every handbook.

Drinkers smoke and smokers drink (8, 9). Identifying and treating smoking, the number-one cause of death among alcoholics, deserves attention (10). Alcohol abusers and alcohol-dependent people often have other substance abuse disorders. Comorbid psychiatric and addiction disorders are the rule rather than the exception (11). Drug testing should be a part of every evaluation of an alcohol abuser or person with any substance abuse disorder. Many patients find it easier to admit to alcohol use than to cocaine or heroin abuse. Denial, minimizing, and outright lying are a part of the disease of alcoholism. Alcohol intoxication diagnoses are made more of-
ten by law enforcement agents using roadside sobriety and breath testing or emergency room personnel using these tests than psychiatrists who only ask the patient. History taking from the patient may be the best example of a clinical oxymoron. Interviews with family members, friends, employers, and others as well as comprehensive testing are essential to avoid either misdiagnosis or partial treatment of disease. The obvious limitations of DSM and psychiatric approaches to diagnosis should be considered alongside the strengths and limitations of drug testing.

Finally, the text should consider an expert’s role in identifying and treating alcoholic colleagues as well as studying and improving treatment. What standards should we follow to evaluate new treatments? Outcome measures such as return to work, psychiatric progress, and drug-free outcomes as demonstrated by 5 years of randomized and supervised testing, the mainstay of physician treatment program evaluations, should be part of all treatment evaluation (4).

In summary, the Handbook of Clinical Alcoholism Treatment is a thorough, well-written, and easy to read resource on a remarkably broad variety of topics within the field of alcoholism. It is a multiauthor or multie xpert text but reads almost as if it were written by the editors. The chapter authors and editors have done a remarkable job in producing a current, comprehensive, extraordinarily informative text for promoting effective intervention and treatment of patients suffering from alcoholism.

References

3. Gold MS, Aronson MD: Treatment of Alcohol Abuse and Dependence. Cambridge, Mass, Harvard University UpToDate, 2004

MARK S. GOLD, M.D.
JOHN F. BRANDT, M.D.
Gainesville, Fla.


This volume is indeed a comprehensive compendium on the subject of substance abuse. I had the privilege several years ago to review the second edition of this volume for the Journal. This third edition is even more complete and more comprehensive.

The editors have brought together more than 90 contributors to this volume. All are knowledgeable about their areas of focus, and they range from individuals whose names are synonymous with substance abuse treatment to those who are on the cutting edge of research for new approaches and new treatments. Despite the large number of contributors, the volume has an internal consistency in approach and presentation that aids the reader and clinician.

This text presents initial sections on the commonalities relative to substances of abuse and the biochemical and physiological responses to these substances. The majority of the text then focuses quite directly on the treatment of each of the major areas of abuse. The initial sections are thus more research and technically oriented and present the reader with the latest in the biological substrates of abuse, and the later chapters are much more practical and specific for the more general reader.

Within each category, the book presents the history of the attempts at treatment for each substance, the characteristics of the abuse syndrome, the different current treatment methodologies and their effectiveness, and the current and future challenges. The book presents the treatment options clearly and in concert with the latest outcome research. It does not simply give credence to “sacred cows,” when, for instance, it notes that “mandatory” meetings of Alcoholics Anonymous as a sole treatment modality is not “proven” to be always effective.

One very small concern is that although there is an expanded section in this volume on the needs and treatments of special populations, the special needs of ethnic minorities and the elderly are still compressed into one excellent but too short chapter. Both of these groups pose very specific and unique problems. With the aging of the general population, substance abuse and even medication confusion in the elderly will be a growing problem, and one that should be given an expanded specific focus in such a volume.

This volume, however, is quite excellent. It is written in such a manner and with enough breadth that it should be read by and on the reference shelves of not only family physicians treating a whole array of substance-abusing patients but also the addictionologist specializing in the latest therapies in the field.

ARON S. WOLF, M.D.
Anchorage, Alaska


These books are volumes two and three of the Key Readings in Addiction Psychiatry series published by the American Academy of Addiction Psychiatry. Each consists of previously published articles that the editors have selected from journals including the American Journal of Addiction, American Journal of Psychiatry, and Archives of General Psychiatry.

Dual Diagnosis begins with a study by Kendler and colleagues of the sources of individual differences in risk for use and abuse of different illicit psychoactive substances in a large group of male twins in order to measure the genetic influence on that risk. Another team evaluated treatment-seeking substance abusers to delineate subgroups of psychiatric comorbidity in terms of DSM-III-R diagnostic criteria in order to facilitate treatment planning. Other researchers addressed the question of whether psychiatric disorders cause or lead to substance abuse or vice versa, with special emphasis on phobias and personality disorders. There is a chapter on the validity and reliability of two instruments used in screening for drug and alcohol abuse. Another explores the rationale for integrated services for dually diagnosed patients. There is a report on the value of one-session motivational interviewing as a prelude to treatment.

One chapter discusses treating depressed alcohol-dependent individuals, and another tries to measure the utility of psychosocial interventions such as psychotherapy and Alcoholics Anonymous (AA) in treating substance-dependent patients with bipolar disorder. Researchers from the Netherlands screened 116 drug-abusing patients for anxiety and mood disorders and compared three means of doing so: a symptom checklist before detoxification, a symptom checklist after detoxification, and a structured interview. A group of Yale psychiatrists review the literature on the association of posttraumatic stress disorder with substance abuse. They cover phenomenology, epidemiology, and neurophysiology. They report promising developments in the pharmacological treatment of this disorder. One article reports a study of correlates of substance abuse in patients with schizophrenia, including intensity of symptoms, demographics, medication side effects, and social functioning. The final chapter reports on a prospective study of the value of personality assessment in predicting vulnerability to substance abuse, with a review of the pertinent literature.

The collection of articles in Psychosocial Treatments focuses on interventions that are supported by research data as well as on their conceptual bases. The first chapter is devoted to the Marlatt-Gordon cognitive behavior model for relapse prevention in alcoholics. It was interesting to read about identifying the different factors that predispose to relapse and the strategies to deal with them, but I was surprised that the authors never mentioned the AA 12-step fellowship, which has been done that since the 1930s. The following chapter explains and discusses motivational interviewing, and the next reports a study of its effectiveness in getting substance abusers into treatment.

Chapter 4 is about network therapy with cocaine abusers, i.e., doing cognitive behavior therapy with peers and relatives concomitantly with the identified patient. Following that is an article on the community-reinforcement approach to treating alcoholism. Its essence is the use of many and varied resources in the patient’s social environment to eliminate positive reinforcement for drinking and maximize positive reinforcement for sobriety. Chapter 6 is a review of the literature on the use of vouchers, i.e., money, to augment other strategies for treating chemical dependencies from nicotine to opiates. The authors acknowledge the difficulty of expanding such an approach from research clinics, where the money comes from grants, to community programs. The following chapter reports a study comparing contingency management with cognitive behavior therapy and a combination of the two strategies in a large group of cocaine-dependent individuals who were in a methadone maintenance treatment program for their opiate addiction.

One chapter reports a survey of self-help strategies among 642 patients with substance abuse disorders, 78% of whom used some form of self-help, such as making efforts to reduce their drug use, changing friends, moving their residence, or joining a self-help group. Another compared two types of residential treatment programs for dually diagnosed veterans—freestanding substance abuse facilities and ones in which psychiatric and substance abuse treatment are integrated. There is a chapter on psychotherapy for substance-abusing adolescents comparing cognitive behavior treatment with interactive treatment. The last chapter explores links between alcohol and crime and offers suggestions for community intervention.

I believe that these 23 articles are well worth reading if, as the editors say, they are carefully chosen samples of the addiction literature. As a general psychiatrist, I may not have seen any of them in the usual course of my reading.

WILLIAM R. FLYNN, M.D.
Napa, Calif.

Cultural Psychiatry


The authors of this revised edition of a 1989 publication are psychologists. The book consists of an introduction and overview and chapters on American Indians, African Americans and racially mixed population groups, Asian Americans, Filipinos, Central Americans, Southeast Asians, and Hispanic-Latino populations. The book is intended to meet the needs of those who are working with the “burgeoning population of racially and culturally diverse youth.”

As a longtime consultant to child guidance clinics and public school systems who serves a large percentage of minority groups, especially African Americans and Hispanics, I read this book with much interest. There are, indeed, major problems facing many children of color and those who serve them,
educators and mental health professionals alike. The authors
opine that culture is “at the heart” of the economic and edu-
cational problems that different minority groups encounter.
They rail against a system which demands that children learn
English and conform their behavior to community standards.
They deplore what they term a “conservative backlash” that
holds parents responsible for their children’s failures. They
feel that “Anglo-Saxons” are responsible for the existence of
“castelike minorities,” specifically, African Americans, Ameri-
can Indians, Mexican Americans, and Puerto Ricans. This
book is written with much emotion, conviction, and intensity.

The authors claim a developmental perspective but omit
discussion of hereditary, gestational, neonatal, psychody-
namic, and other early childhood influences and factors in
the formation of adaptive capacities. Recognition of individ-
ual differences in talents and abilities is deemed irrelevant be-
cause of what the authors regard as invalid or unproven tests
and measurements. However, when evaluating children and
adolescents of any color, I often find it very helpful to con-
sider the results of tests such as the WISC, Behavior Assess-
ment System for Children, and MMPI-A in arriving at diag-
noses and recommendations.

The authors do not examine how it is that many children
and adolescents of every color succeed, nor do they attempt
to account for why members of some ethnic groups do better
on average than members of other groups. I am reminded, for
example, of “Louie,” a Mexican boy from Saginaw, Mich.,
whose family lived in the city’s poorest area. Louie’s father was
a postal clerk, and he had five brothers and sisters, all of
whom were outstanding students. Louie was an excellent musi-
cian, athlete, and student. He finished at the top of his large
public high school class and eventually graduated from one of
the nation’s best medical schools. He was smart, hardworking,
and made no excuses.

Unfortunately, these authors blame what they regard as the
ruling majority for the disappointments too many children of
color experience in their pursuit of their interpretation of the
American dream. Accordingly, they argue for allocation of
more money from the government to level the cultural play-
ning field. The authors are four-square against English-only
programs and opine that the

Clinician’s Guide to Cultural Psychiatry, by Wen-Shing
$59.95 (paper).

For several decades, Dr. Tseng has studied “the unique be-
behavior patterns and life style shared by a group of people,
which distinguish it from others.” The result, Handbook of
Cultural Psychiatry (1), was published in 2001. The current
compendium derives largely from it.

Clinician’s Guide to Cultural Psychiatry starts with chapters
that define cultural issues and stress reactions. These chap-
ters could be considered the theoretical basis for the book, re-
affirming that most pathological manifestations are influ-
enced by the individual’s cultural group and his or her
environment.

Chapters 3 and 4 focus on specific phenomena and general
disorders. The specific phenomena addressed in chapter 3 in-
clude the many specific culture-related syndromes that in the
past constituted a great part of “cross-cultural psychiatry.”
Chapter 4 refers to the psychiatric disorders usually ad-
dressed in the psychiatric nomenclature. The author reminds
us that the Collaborative Study of Psychological Problems in
General Health Care, a large-scale international multisite
comparative investigation carried out by the World Health
Organization, showed striking similarities in the distribution
of a number of disorders around the world. Much more re-
search is necessary to appreciate whether cultural factors
make a difference in the presentation and diagnosis of most
disorders.

The remaining chapters are more relevant to clinicians.
How do we take cultural factors into account in the diagnostic
assessment, psychotherapy, drug therapy, and choice of treat-
ment settings? A good beginning might be to accept that cul-

ual differences really exist and may facilitate or interfere
with the work of the clinician; that psychotherapy may be
more effective if congruent with the patient’s culture and val-
ues; that biological factors enhance or reduce the effects of
medications; that culture may be a major factor in accepting
a treatment setting.

The clinicians who may fish in this well-stocked sea of
clinically relevant issues will be disappointed with their
catch. They may not be satisfied with the size of the offerings
(four pages on Hispanic American groups, one page on cog-
nitive therapy). They may be happier if they see the book
more as a display of potentially relevant issues than as a
manual for the clinician eager to learn practical pointers on cultural psychiatry.

Reference

RODRIGO A. MUÑOZ, M.D.
San Diego, Calif.

Reprints are not available; however, Book Forum reviews can be downloaded at http://ajp.psychiatryonline.org.

Corrections
The first author of the article “Gamma Synchrony’ in First-Episode Schizophrenia: A Disorder of Temporal Connectivity?” (Am J Psychiatry 2005; 162:459–465) should have been listed as Matthew P. Symond, D.C.P., M.Sc.”

In a letter to the editor in the March issue by Christopher Thomas, Pharm.D., B.C.P.P., titled “Memantine and Catatonic Schizophrenia” (Am J. Psychiatry 2005; 162:826), additional authors should have been listed. They are as follows: Brendan T. Carroll, M.D., Robert T. Maley, Pharm.D., Kameshwari Jayanti, M.D., and Atchuthamba Koduri, M.D.