JOURNAL OF CHILD AND ADOLESCENT PSYCHOPHARMACOLOGY
Volume 13, Number 1, 2003

Pp. 1–15

Methodological Issues and Controversies in Clinical Trials with Child and Adolescent Patients with Bipolar Disorder: Report of a Consensus Conference[1]

GABRIELLE A. CARLSON, MD, PETER S. JENSEN, MD, ROBERT L. FINDLING, MD, ROGER E. MEYER, MD, JOSEPH CALABRESE, MD, MELISSA P. DELBELLO, MD, GRAHAM EMSLIE, MD, LAURIE FLYNN, FREDERICK GOODWIN, MD, MARTHA HELLANDER, ROBERT KOWATCH, MD, VIVEK KUSUMAKAR, MD, THOMAS LAUGHREN, MD[1], ELLEN LEIBENLUFT, MD, JAMES MCCRACKEN, MD, EDITHA NOTTELMANN, PhD, DANIEL PINE, MD, GARY SACHS, MD, DAVID SHAFFER, MD, RENEE SIMAR, PhD, MICHAEL STROBER, PhD, ELIZABETH B. WELLER, MD, JANET WOZNIAK, MD, and ERIC A. YOUNGSTROM, PhD

Abstract

Objective: To achieve consensus among researchers, pharmaceutical industry representatives, federal regulatory agency staff, and family advocates on a template for clinical trials of acute mania/bipolar disorder in children and adolescents.

Method: The American Academy of Child and Adolescent Psychiatry, in collaboration with Best Practice, convened a group of experts from the key stakeholder communities (including adult psychiatrists with expertise in bipolar disorder) and assigned them to workgroups to examine core methodological issues surrounding the design of clinical trials and, ultimately, to generate a consensus statement encompassing: (1) inclusion/exclusion criteria, (2) investigator training needs and site selection, (3) assessment and outcome measures, (4) protocol design and ethical issues unique to trials involving children/adolescents, and (5) regulatory agency perspectives on these deliberations.

Best Practice, Bethesda, Maryland (Drs. Meyer and Goodwin); Case Western Reserve University, Cleveland, Ohio (Drs. Calabrese, Findling, and Youngstrom); Child and Adolescent Bipolar Foundation, Wilmette, Illinois (Ms. Hellander); Columbia University, New York, New York (Drs. Jensen and Shaffer and Ms. Flynn); Dalhousie University, Halifax, Nova Scotia, Canada (Dr. Kusumakar); Food and Drug Administration, Rockville, Maryland (Dr. Laughren); George Washington University, Washington, DC (Dr. Goodwin); Harvard University, Boston, Massachusetts (Drs. Sachs and Wozniak); INC Research, Austin, Texas (Dr. Simar); National Institute of Mental Health, Bethesda, Maryland (Drs. Liebenluft, Nottelmann, and Pine); Stony Brook University, Stony Brook, New York (Dr. Carlson); University of California-Los Angeles (Drs. McCracken and Strober); University of Cincinnati, Cincinnati, Ohio (Drs. DelBello and Kowatch); University of Pennsylvania-Children's Hospital of Philadelphia (Dr. Weller); University of Texas Southwestern Medical Center, Dallas, Texas (Dr. Emslie).

[1]Dr. Laughren is with the Food and Drug Administration, Rockville, Maryland. However, Dr. Laughren's contribution to this article was made in his private capacity. No official support or endorsement by the Food and Drug Administration is intended or should be inferred.

[1]This conference was convened by the American Academy of Child and Adolescent Psychiatry and funded by Best Practice with partial financial support from unrestricted grants from the following companies: Abbott Pharmaceuticals, Bristol-Myers Squibb, GlaxoSmithKline, INC Research, Janssen Pharmaceutica, Johnson & Johnson Pharmaceutical Research and Development, Eli Lilly, Novartis, Pfizer, and Solvay Pharmaceuticals.

Results: Conference participants reached agreement on 18 broad methodological questions. Key points of consensus were to assign priority to placebo-controlled studies of acute manic episodes in children and adolescents aged 10–17 years, who may or may not be hospitalized, and who may or may not suffer from common comorbid psychiatric disorders; to require that specialist diagnostic "gatekeepers" screen youths' eligibility to participate in trials; to monitor interviewer and rater competency over the course of the trial using agreed upon standards; and to develop new tools for assessment, including scales to measure aggression/ rage and cognitive function, while using the best available instruments (e.g., Young Mania Rating Scale) in the interim.

Conclusions: Methodologically rigorous, large-scale clinical trials of treatment of acute mania are urgently needed to provide information regarding the safety and efficacy, in youth, of diverse agents with potential mood-stabilizing properties.

Key Words: mania, bipolar spectrum, bipolar–not otherwise specified (BP-NOS), clinical trials, assessment, ethics.

Introduction

Several trends recently have converged to add urgency to the need to address unresolved issues surrounding the conduct of psychopharmacologic treatment studies in child/adolescent bipolar disorder. These trends include the increasing frequency with which bipolar disorder is being diagnosed in children and adolescents, an increase in off-label prescribing of antipsychotic and anticonvulsant medications to treat children/adolescents presumed to have the disorder, and mounting pressure on the pharmaceutical industry to include in clinical trials the specific populations for whom medications will be prescribed. However, considerable controversy persists surrounding core issues that are central to successful clinical trials of bipolar disorder in younger populations, including, for example, the definition of the illness in this age group and measures to assess it (Biederman et al., 1998; Carlson, 1998, 1999). In the absence of consensus among researchers, industry sponsors, federal regulators, and families on both the feasibility of and ground rules for conducting high-quality clinical trials for this population, children and adolescents will likely be deprived of the opportunity to receive efficacious and effective treatments for bipolar disorder, and the field will remain at a standstill.

In response to these concerns, the American Academy of Child and Adolescent Psychiatry, in collaboration with Best Practice, convened in June 2002 a working conference on "Methodological Issues and Controversies in Clinical Trials with Child and Adolescent Patients with Bipolar Disorder." The explicit purpose of the meeting was to develop a template for clinical trials of acute mania/bipolar disorder in children and adolescents. Invited participants included clinical researchers with expertise in childhood and adult bipolar illness, pharmaceutical industry sponsors with an interest in mood stabilizer products, staff of the Food and Drug Administration (FDA; and their counterparts from regulatory agencies in Canada and the European Union) and the National Institute of Mental Health, and representatives of families with affected children. Participants are listed at the end of the text.

The conference opened with plenary presentations designed to articulate, from the vantage points of the research community, industry, and regulatory agencies, the current status of needs, barriers, and options in the design of clinical trials for children and adolescents with mania/bipolar spectrum illness. Following the plenary presentations, all participants were assigned to breakout workgroups and charged to examine specific questions and seek consensus on controversial issues that had been identified prior to the meeting by the organizers (Gabrielle Carlson, Robert Findling, Peter Jensen, and Roger Meyer). Four workgroups respectively addressed the following topics: (1) Who should be studied (inclusion/exclusion criteria)? (2) Who should conduct trials (training and site selection)? (3) What assessment and outcome measures are available, appropriate, and needed? (4) What protocol design and ethical concerns are unique to clinical trials involving children and adolescents with bipolar disorder? The workgroups reconvened in full session twice, allowing all forum participants the opportunity to hear and comment on other workgroups' progress toward consensus. In addition, on the second day, a fifth workgroup met to enable representatives of the FDA as well as European and Canadian regulatory bodies to discuss recommendations emerging from Workgroups I–IV. A Co-Chair from each of the four other workgroups joined Workgroup V[ii].

This brief report presents the consensus recommendations and summarizes the discussions surrounding them. Questions that guided the discussions have been clustered in the text under the relevant workgroup's section of this

[ii]The other four workgroups continued to discuss their issues and brought their conclusions to the final plenary session at the end of the second day of deliberations.

article. A list of future research needs is presented at the end of the article.

Workgroup I: Who Should Be Studied?

Workgroup I addressed inclusion and exclusion criteria for clinical trials of child/adolescent bipolar disorder with respect to the age range of subjects, subtype of the disorder, comorbidity, and clinical severity.

Consensus Recommendations:

- Children and adolescents with acute mania should be given immediate priority for bipolar treatment studies. Inclusion criteria for these treatment studies should include acute mania in patients (age 10–17 years) who meet DSM-IV American Psychiatric Association (APA 1994) criteria for Bipolar I disorder.
- Treatment studies may include inpatients, outpatients, and/or day patients, based on clinical judgment. For outpatients and day patients, however, safety concerns must be carefully addressed with good clinical practices such as 24-hour access to emergency care including (if indicated) inpatient stabilization, removal of firearms and sharps, monitoring of suicide risk, parent education about the disorder and its management, and family support services. Future studies should determine whether the inclusion of outpatients leads to poor drug-placebo differentiation.
- Children/adolescents with comorbid attention deficit hyperactivity disorder (ADHD), conduct disorder, or substance abuse should not be excluded a priori from mania treatment studies. However, children with autism, IQ < 70, substance-induced mania, unstable neurological conditions, substance dependence, and serious homicidal/suicidal tendencies should be excluded.

Discussion:

Workgroup I addressed the question of who should be included in industry-supported trials of pediatric/adolescent bipolar disorder. When industry proposes extending an adult indication to subjects younger than age 18, the FDA, through a written request, specifies what would be needed in a development program to support a claim for this indication (e.g., bipolar mania in children and adolescents). The current approach, as illustrated in written requests that have already been issued, is to target acute mania in a pediatric age group in which the manic syndrome is phenomenologically similar to that seen in adults.

In the United States, studies of adult bipolar disorder have concentrated mainly on "acute mania," Bipolar I, with hospitalized samples, although outpatient augmentation trials have been done as well. In Europe, according to the participant from the European Union, efficacy must be demonstrated in short-term studies that show an effect in manic episodes, and it should be shown that efficacy is to be maintained during the episode (Committee for Proprietary Medicinal Products [CPMP], 2001); he noted, furthermore, that long-term safety studies with regard to physical, sexual (endocrinological), and mental development are needed. The FDA has not specified how long the subject needs to have been manic, but the need for hospitalization has provided some measure of severity if the episode has been shorter than 7 days. Given this experience with adult bipolar disorder, and the FDA requirements as reflected in a template written request provided to forum participants, the discussion on inclusion and exclusion criteria for children and adolescents with bipolar disorder developed as follows:

Inclusion criteria: Child/adolescent bipolar treatment studies must specify acute mania (i.e., patients who meet DSM-IV criteria for Bipolar I disorder in a current manic episode). This requires: a "distinct period" of abnormally and persistently elevated, expansive, or irritable mood lasting at least 1 week, accompanied by at least three of the following symptoms (four, if irritability is the only mood symptom): significant levels of inflated self-esteem or grandiosity, decreased need for sleep, greater talkativeness than usual (or pressure to keep talking), flight of ideas or subjective experience of racing thoughts, distractibility, increase in goal-directed activity or psychomotor agitation, and excessive involvement in pleasurable activities that have a high potential for painful consequences (American Psychiatric Association, 1994). These symptoms must be markedly impairing and not caused by a medical condition, medication, or illicit drug. Duration of less than 7 days is acceptable if hospitalization for the manic episode is necessary. Finally, the manic phase must represent a change from previous functioning. Forum participants acknowledged that shorter episodes and rapid cycles commonly observed in children and adolescents may in fact be harbingers of episodes meeting full DSM-IV duration criteria and may warrant further definition and study. However, altering the DSM criteria to make exceptions for children is not an acceptable option at this point. The definition of what constitutes an episode (including duration and severity) remained controversial. The group resolved that it will be critical to develop more rigorous and specific criteria for bipolar—not otherwise specified (BP-NOS), which may be a more common presentation in children and adolescents than distinct episodes of acute mania, before these patients should be included in large-scale, industry-sponsored trials.

Although the most recent written request for a clinical trial for treatment of acute mania in young people was limited to adolescents, the forum recommended age 10 years as the lower age limit at which sufficient numbers

of patients with adult-type mania can be located and recruited to participate in industry-supported trials. Participants agreed that acute mania with clear episodes can be found in children younger than 10 years, but they did not reach consensus on how frequently it occurs or how readily professionals lacking extensive clinical experience could recognize and distinguish it from other conditions. Without appropriate attention to the influence of comorbidity and developmental stage on such factors as cognition, concreteness of language, activity level, and capacity for emotional regulation, it is possible to mistake normal childhood utterances or behaviors for mania or to misinterpret the symptoms and behaviors of other conditions as symptoms of mania. Moreover, representatives of regulatory agencies suggested that lowering the age limit to 7 or 8 years of age would oblige the sponsor to draw 40% of the sample from children 8–12 years of age. Given the uncertainties regarding the prevalence of mania in children younger than 10 years, and the limited availability of necessary diagnostic expertise, lowering the age for inclusion in the study would make it logistically difficult in most situations.

Inpatient versus outpatient: The inpatient setting for studies of acute mania in adults has been used for several reasons. The first is to ensure sufficient severity to obtain drug/placebo differences if they exist. When outpatients have been studied, the placebo response rate has increased considerably. The second reason is to ensure patient safety, particularly in the placebo condition. The third is to ensure adherence with regular medication administration at a time when the patient's condition severely limits his or her judgment. Studies in adult inpatients with acute mania also reflect the reality that the vast majority of adults with this diagnosis are hospitalized. In contrast, the vast majority of children and adolescents with this diagnosis are treated as outpatients or partial hospital patients, not in inpatient units. For this reason, forum participants concluded that restricting a study population of young people to inpatients only was impractical, unnecessarily restrictive, and would yield results that would not generalize to the population of children and adolescents presenting with acute mania. This consideration led to agreement that inpatient, outpatient, and/or day hospital patients could be included based on clinical judgment and parental agreement. As noted, inclusion of children from outpatient/day treatment sources in placebo-controlled treatment studies underscores the need for special effort to address safety concerns. Suggested protective steps include 24-hour access to emergency care, including, if indicated, inpatient stabilization, removal of firearms/sharps from the home/outpatient environment, educating the child's caretakers about suicide risk and strategies for responding to suicidal behavior, and educating parents and children about how to manage the illness. Other interventions in lieu of hospitalization might include respite workers or in-home, one-to-one aides to families in both arms of a placebo study.

Exclusionary criteria: Exclusionary criteria are used to minimize variance within a study population so as to increase statistical power and diminish the likelihood of study failure. Standard exclusionary criteria in adult trials are serious psychiatric (e.g., substance dependence, eating disorders), medical, and neurological comorbidity. In children and adolescents, the comorbidities are not only somewhat different but are so prevalent that confining samples to those without any comorbidity would be fruitless. Participants readily agreed on the need to exclude patients with schizophrenia, autism, unstable medical and neurological conditions that would interfere with study compliance, substance/medication-induced mania, or patients with IQ's less than 70, but discussed at length the eligibility of children with comorbid ADHD and suicidal ideation. The group concluded that given the high prevalence of ADHD comorbid with presumed and documented Bipolar I disorder, children with ADHD should not be routinely excluded from treatment studies for bipolar disorder. When such children are included in a BP-I clinical trial, stimulants should be discontinued for an agreed upon washout period and throughout the randomization period of an acute (i.e., 21-day) clinical trial.

One of the participants articulating the views of the CPMP, the group responsible for advising the European Commission as to whether or not medicinal products should be granted a marketing authorization in the European Union, observed that, in principle, to grant a license in Europe, separate studies in children and adolescents are needed (CPMP, 2000). In Europe, there is ongoing discussion about the validity of manic episodes in children, and specific guidelines should wait until this discussion is settled. He also opined, apart from his CPMP role, that by including children/adolescents with comorbid ADHD, it will not be possible to differentiate between symptom reduction in manic symptoms and ADHD symptoms. Thus, he suggested that to demonstrate efficacy in the manic episode unequivocally, clinically relevant comorbid disorders should be excluded.

Participants initially listed suicidal ideation, along with risk of dangerousness to others, among the exclusion criteria, but discussion regarding operationalizing procedures to ensure a safe environment for children and adolescents in outpatient treatment studies suggested that suicidal ideation need not be an automatic cause for exclusion. Rather, each prospective participant should receive a thorough evaluation for suicidal risk, including the presence of a plan, a history of prior suicide attempts, a history of suicide in a family member or close friend, and a determination whether the subject will agree to contact the investigator immediately in case of suicidal thoughts. If the investigator

determines that the degree of suicide risk can be safely managed, a potential subject might be enrolled in a trial. In such cases, the investigator and the research participant (and his/her family) should have access to 24-hour, 7-day/week emergency care as needed (see below).

Adult trials typically exclude "first episode" patients and patients who have failed multiple treatments. In the first instance, the placebo response rate is very high. In the second, drug response is very low. Forum participants agreed, however, that excluding first episode children and/or adolescent patients would be unnecessarily restrictive. They further agreed that if a patient had been exposed to a medication clinically for an adequate period (e.g., 6 weeks) without responding positively, it would not be appropriate to retry them on that same medication in a clinical trial. That is, rather than excluding subjects simply on the basis of their having failed to respond to antimanic medications in the past, only subjects who have failed to respond to or were unable to tolerate a prior adequate dosing trial of the specific medication under study should be excluded.

Workgroup II: Who Should Do the Studying?

There are relative paucities both of investigators with the expertise needed to conduct rigorous clinical trials of children/adolescents with bipolar mania and of young subjects who meet DSM-IV criteria for this condition. Accordingly, Workgroup II addressed questions regarding the nature of the initial assessment, including the qualifications and credentials of the lead clinician/diagnostician, instruments used, the identity of informants other than the child, training requirements for all members of the clinical trial research team, and strategies for ensuring that all sites in a multisite trial have the necessary expertise and the capacity to recruit a sufficient number of subjects to permit adequately powered analyses.

Consensus Recommendations:

- Bipolar diagnoses should be established by a well-trained (doctoral level) child mental health specialist (psychologist or psychiatrist).
- Diagnoses should be based on a minimum of two sources of information (e.g., patient, mother, father, teacher, clergy), across two settings (e.g., home, school, camp, house of worship, athletic leagues), and/or with direct observation.
- A separate site should monitor recruitment and diagnostic assessments for both the sponsor and contract research organization. A contract research organization provides a variety of services to industry sponsors, including phase I study sites, laboratory services, ECG monitoring, pharmacokinetics, protocol design, FDA regulatory advice, data management, data analysis, study monitoring, and other services. Monitoring (whether by the sponsor or contract research organization) should include a review of a random subset of diagnostic videotapes. If younger aged children are included, their tapes should be oversampled. Cross-site validational studies would be helpful.
- Objective measures and standards for interviewer and rater competency should be established, including specification of a minimum number of hours of training.
- New tools should be developed to assist the field; these should include, but not be limited to, manualization of the Young Mania Rating Scale (Y-MRS; Young et al., 1978) and videotape libraries.

Discussion:

Assessment: In adults, mania is such a sufficiently remarkable and common reason for admission to inpatient services that recognizing the condition and distinguishing it from schizophrenia and other disorders is part of the day-to-day experience of adult psychiatrists. In efficacy trials, patient selection is ultimately systematized by the use of structured interviews, most commonly the Structured Clinical Interview for DSM-IV (SCID-I; First et al., 1995). Consent and symptom information is obtained from the manic subject, and although experienced interviewers tend to prefer to obtain this with a "significant other" informant present, a second source is not required. The patient needs to be sufficiently symptomatic that diagnosis is obvious or sufficiently articulate to describe history and symptoms of impairment. Studies require that a patient must meet both diagnostic and severity criteria, with severity gauged either by the Y-MRS (Young et al., 1978) or the Schedule for Affective Disorders and Schizophrenia—Current (SADS-C; Endicott and Spitzer, 1978) at two time points, typically at the initial assessment and at baseline for entry into a trial.

Assessment of child/adolescent bipolar disorder poses a different set of considerations. First, there is a lack of agreement/clarity as to the clinical entity of childhood bipolar disorder. The differential diagnosis (ADHD, pervasive developmental disorder) differs from that seen in adult bipolar disorder, and the impact of development on ascertainment and assessment of symptoms adds special challenges that necessitate modifications of standard adult assessment/diagnostic procedures. Tailoring these procedures becomes more important when younger children are included in a study, but it is necessary even for older children and young adolescents. For that reason, forum participants agreed that assessment/diagnosis should begin with a thorough assessment conducted by a well-

trained child/adolescent psychiatrist or psychologist. The clinician must obtain information from multiple informants (parent/caretaker and child, as well as possibly siblings, teachers, or others) and ascertain the presence of symptoms in at least two settings (home plus another) to confirm symptoms and reinforce the credibility of the diagnosis.

During the assessment process, the interview with the child should focus on internalizing symptoms (i.e., the subjective experience of mania, depression, and suicidality), psychosis, thought/language disorder, and pervasive developmental disorder, whereas interviews with parent(s) or other informants should focus on externalizing symptoms (i.e., observable symptoms of mania) as well as the usual panoply of psychopathology. As with adults, initial screening and baseline ratings should be separated in time and should use the same informants to help ensure reliability. Although it is not necessary to have all informants agree on all symptoms in every setting, forum participants strongly agreed that an assessment based on input from multiple informants commenting on a child's behavior *in more than one setting* would be most likely to ensure a valid diagnosis and reliable assessments of impairment and severity. Several participants emphasized the importance of teacher ratings. Although demands on teachers' time often make these ratings difficult to obtain, having a picture of a child's behavior in school is critical, especially if the patient is unable to articulate his/ her symptoms clearly. In teens, the guidance counselor may be helpful.

Workgroup members underscored the importance of accurately understanding a child's mental age during the assessment process and noted the need for employing developmental considerations in any diagnostic scheme and assessment. The clinical interview, which initially selects patients for trials, needs to be administered by a well-trained and experienced doctoral-level mental health clinician who serves as a gate-keeper to ensure that the right subjects are selected. The necessary qualification of the person doing the subsequent interview, which quantifies and confirms clinical information, depends on whether it is structured/respondent-based **or semi-structured/interviewer driven.** Examples of structured interviews include the Diagnostic Interview Schedule for Children (DISC; Shaffer et al., 1996), the Diagnostic Interview for Children and Adolescents (DICA; Herjanic and Reich, 1977, 1982), and the Schedule for Affective Disorders and Schizophrenia for Children (K-SADS; Chambers et al., 1985; Kaufman et al., 1997). The latter is a semistructured interview that affords a clinician the flexibility to consider issues that may not be articulated during a structured interview. The primary disadvantage of a semistructured interview is that it requires extensive training to be used properly. Respondent-based structured interviews afford less disclosure but effectively pace an interview, reduce input errors and need for editing, and facilitate data management. Although a structured interview is useful for identifying the diverse symptoms that may be seen in a variety of disorders, it does not afford the opportunity to identify questions that may have been misunderstood by the respondent

Forum participants did not reach consensus on whether a structured or semistructured interview was absolutely necessary in the assessment but agreed that if either is used, the results of the interview should be consistent with the clinical diagnostic impression of the child psychiatrist/ psychologist. Noting that the K-SADS interview requires that the interviewer have an accurate, thorough understanding of the condition being investigated, several discussants voiced concern that it was being used more like a structured/respondent-based interview but without the requisite structure. Although many participants felt that all pertinent modules of the K-SADS should be used (i.e., not just the mood disorders module), others cautioned that the time necessary to do this represented a significant disincentive.

Other interviewer-based tools such as the Hamilton Depression Rating Scale (Hamilton, 1960), the Clinical Global Impressions Scale (Guy, 1976), the Y-MRS (Young et al., 1978), and the Childhood Depression Rating Scale-Revised (CDRS-R; Poznanski et al., 1983) are not diagnostic instruments. Rather, they offer both a means of standardizing a clinician's impressions and a critical threshold of symptom severity. Participants agreed that ongoing ratings with these instruments could be performed by adequately trained BA- and MA-level research personnel. Discussion also addressed the importance of defining the bipolar syndrome consistently across settings and studies as well as using the same rating scales across settings and studies. Although acute mania was the target condition for treatment in the clinical trials that were the subject of the forum, the group emphasized the need to know the mix of Bipolar I and II disorders and BP-NOS as seen in subjects in validation samples done with any of the different versions of various instruments.

Site selection: Many sites will be needed in the near future to study the efficacy and safety of medication treatments for mania in children. Although a smaller number of sites confers the advantage of improved reliability of diagnoses and measurements, limited numbers of patients are available at sites possessing the requisite research expertise. An examination of treatment effect size in one open-label study suggests that it may be necessary to enroll approximately 100 patients for each treatment under investigation (Kowatch et al., 2000). It is essential to educate the investigators and raters at nonacademic and less-research-experienced academic sites about the proper use of assessment/rating instruments and the importance of using them consistently with young subjects and their parents. Forum participants discussed at length the need for objective measures of rater competency, including a required minimum number of hours of training. Partici-

pants did not feel that investigator meetings were adequate for addressing all of the training issues associated with a large clinical trial with this subject population.

Various strategies for ensuring quality control were considered, including (1) having experienced researchers monitor videotaped interviews, either comprehensively or randomly; (2) developing web-based training, which might be useful for initiating new sites or new raters as the trial continues; (3) assigning responsibility to one site with well-trained personnel for monitoring recruitment at all participating sites; and/or (4) manualizing key rating instruments. Training modules could utilize videotape or DVD libraries of interviews with parents and the affected child, as well as observational tapes of the child in a natural setting (waiting room, at play, etc.), plus a casebook containing family and educational history, psychometric testing results, and other relevant materials. Finally, participants recommended that industry sponsors support efforts to develop guidelines for training and for certification of sites capable of conducting rigorous clinical trials with children and adolescents.

Conducting a clinical trial requires the collaboration of three parties: the industry sponsor, the patient participant, and the specific site executing the measures and treatments. Each participant must feel invested in the endeavor. Research sites assume a significant care burden of managing a group of young people with mania of sufficient severity that they warrant hospitalization, even if the trial protocol does not require hospitalization. In addition, institutional review board requirements for clinical research are becoming increasingly complex and time-consuming. Participants agreed that data sharing and timely publication, including publication of negative studies, are important to their willingness to participate in a clinical trial. Finally, conferees noted that careful attention must be paid to cost offset for the clinical care necessary to attract patients to protocols, to meet the ethical requirement of caring for them for an appropriate period of time after termination of the trial, and to start-up and shut-down expenses.

Because of limited clinical trials experience with this patient population outside of academia, it will be important to engage the relatively few academic centers that have such experience in these trials. No single best solution to the problems of site selection emerged. Among the suggestions for recruiting sites were:

- Adopt for use in child psychiatry the "hub and spoke" model used in pediatric oncology to facilitate linkage of community sites with academic centers of research/clinical excellence (Heinig et al., 1999). This model is reported to be particularly effective when a trial is studying a new medication not yet available to community clinicians.
- Encourage sites experienced in conducting clinical trials with adult bipolar patients to "add on" child psychiatric capabilities. Child psychiatry programs could then "borrow" needed research infrastructure elements (e.g., study coordinators) needed to support child psychiatric studies.
- Establish consultancy relationships between academic child psychiatry sites and private contract research organizations.
- Use international sites. Although Europeans/Canadians view bipolar disorder in prepubertal children skeptically, they should be able to conduct trials with older children and adolescents.[iii]
- One or more industry sponsors should consider developing model-collaborative agreements with a critical number of academic sites in child psychiatry in order to ensure adequate numbers of clinical trial sites and investigators. Models of industry/academic collaborations have been developed in other areas (Zisson, 2001). The need is critical throughout child psychiatry but especially so in relation to childhood bipolar disorder.

Workgroup III: Outcome Measures

The success of rigorous clinical trials in child/adolescent populations is contingent, in large part, on the availability of effective assessment and outcome measures that have been designed for and/or validated in the target population. Workgroup III reviewed existing measures and discussed their reliability/validity in younger populations; considered strategies for ascertaining the validity for youth of rating scales that are used successfully in adult trials; addressed the question of most useful secondary outcome measures; and considered the challenge of reconciling information obtained from two or more sources, in addition to the clinical investigator and child/adolescent subject.

Consensus Recommendations:

- The Y-MRS has acceptable reliability and limited validity data for rating mania in children and adolescents and is "good enough" to use as a primary outcome measure in the absence of a better alternative. For rating depressive symptoms in children, the CDRS-Revised is recommended (Poznanski et al., 1985).
- Secondary outcome assessments should include, but not necessarily be limited to, ADHD symptomatology, aggressive behavior, cognitive function, clinical global

[iii]However, multinational trials would need to comply with global and local laws and regulations, and it may not be possible to conduct an adequate and well-controlled trial on pediatric bipolar disorder in some countries due to restrictions on the use of placebo. In any event, site selection would need to be done very carefully, and training criteria would assume heightened significance.

improvement, and, quality of life (including family, school, and peer relationships).
- Additional/better scales will need to be developed to measure irritability, aggression/rage (including the frequency, intensity, and duration of rages), and cognitive function in this population.
- Information obtained from various sources should be gathered, carefully considered, and finally reconciled using the skilled clinical interviewer's best judgment, giving emphasis to validity rather than relying on simple "and/or" rules.

Discussion:

Selecting the correct outcome measure to detect drug response and difference from placebo is as important as selecting the right subjects to participate in a clinical trial. The primary outcome measure (i.e., the one on which the success of the study hinges) must measure mania. In adult studies, secondary outcome measures have addressed clinical global improvement, general psychopathology, psychosis, depression, and functional status. Although several clinician-rated mania scales have been developed over the past 30 years (see Table 1), the Y-MRS (Young et al., 1978) is most frequently used in clinical trials involving adult inpatients with mania. The scale was originally designed to evaluate changes in symptom severity in adult patients hospitalized for mania. It is adequately sensitive to changes in symptom severity among moderately severe or markedly impaired subjects, but it is less sensitive to symptom changes in patients with mild illness severity. Its contemporary widespread use is puzzling in light of the fact that it does not comprehensively cover the DSM-IV criteria for mania. Less commonly used, the SADS-C Mania Rating Scale (Endicott and Spitzer, 1978) takes items from the SADS-Lifetime (Endicott and Spitzer, 1978) interview section on mania to compose a rating scale that has been used in at least one pivotal adult study (Bowden et al., 1994). Finally, many children and adolescents with a bipolar disorder present in a mixed state, and it will be important in the future to develop an instrument that assesses both the manic and depressed symptoms during mixed states.

Some secondary measures commonly used in studies of adult mania, such as the Clinical Global Impressions Scale (Guy, 1976), Brief Psychiatric Rating Scale (BPRS), and the Positive and Negative Symptom Scale (PANSS; Kay et al., 1987), also have been geared to the assessment of severe psychopathology in hospitalized patients. These interview-based measures are used frequently because they appear to be psychometrically acceptable and are sensitive to change (American Psychiatric Association, 2000). Commonly used interview-based depression severity measures include one of many forms of the Hamilton Rating Scale for Depression (Hamilton, 1960) or the Montgomery-Asberg Depression Rating Scale (Asberg et al., 1978; Montgomery and Asberg, 1979). Although the measures of global functioning as well as psychotic and depressive symptomatology can serve to anchor symptom severity, they do not distinguish symptom severity from symptom frequency and/or duration. In inpatient trials, where patients are under frequent observation, this is less of a problem than in outpatient trials (American Psychiatric Association, 2000, p. 533).

Of the measures most commonly used to assess adult bipolar disorder in participants in clinical trials, data regarding their use with children and adolescents exist only for a few, including the Y-MRS for mania and the CDRS-R for depressive symptoms. The BPRS-Children's version (Hughes et al., 2001), which is not related to the adult BPRS, recently has been re-anchored to improve reliability and validity for both trained and untrained raters. Studies underway in 2002 are using the adult PANSS and the BPRS (R.L. Findling, personal communication, June 17, 2002). As previously noted, however, the use of adult measures with children is problematic for several reasons, including the likelihood that the presentation of the disorder will be different at younger ages and the fact that some symptoms being measured may be inappropriate at younger developmental stages. Although the Hamilton Depression Rating Scale has been used for depression studies in adolescents, it has not been shown to be as useful or as sensitive to change as the CDRS-R (Emslie et al., 1997; Poznanski et al., 1983, 1984). Derived from the Hamilton Depression Rating Scale, the latter incorporates symptoms and behaviors that have been ob-

TABLE 1. MANIA RATING SCALES

Rating Scale	*Study*
Beigel-Murphy Manic State Rating Scale	Beigel and Murphy (1971)
The Manic State Rating Scale	Blackburn et al. (1977)
Young Mania Rating Scale	Young et al. (1978)
SADS-C Mania Rating Scale	Endicott and Spitzer (1978)
Bech-Rafaelsen Mania Scale	Bech et al. (1979)
Manchester Nurse Rating Scale for Mania	Brierley et al. (1988)
Clinician-Administered Rating Scale for Mania	Altman et al. (1994)

SADS-C = Schedule for Affective Disorders and Schizophrenia—Current.

served in depressed children and adolescents, and it takes advantage of the likelihood that the investigator may obtain information from several sources.

The major problem confronting conference participants in selecting appropriate outcome measures was that, to date, there have been no double-blind, placebo-controlled studies of acute mania in hospitalized children/ adolescents, and, as a result, no primary outcome measures have been demonstrated to be unequivocally successful. The Clinical Global Impressions Scale (Guy, 1976) and the Global Clinical Judgments Scale (Campbell et al., 1984) have been used, respectively, in an outpatient study and an inpatient discontinuation study of youth mania (V. Kafantaris, personal communication, June 17, 2002). The Y-MRS has been used previously in an open comparison trial of three mood stabilizers (Kowatch et al., 2000) as well as in inpatient, open, and discontinuation studies, and it is being used currently (2002) in several treatment trials (Kowatch, personal communication, June 17, 2002). Although it appears to have adequate reliability, age is a strong covariate in ratings of symptom severity, with younger children likely to be given higher ratings. The Y-MRS recently has been reanchored for younger subjects by a group consensus, but no data were available at the time of the conference on the revised version. The K-SADS Mania Rating Scale (D. Axelson, personal communication, June 17, 2002) appears promising. It functions like the SADS-C, as it is derived from the K-SADS mania section, but has yet to be validated in children and adolescents.

Forum participants agreed that, to advance the field, rating scales that have been used successfully in adult trials should be used concurrently with those that are being developed; for example, the Y-MRS should be administered with the K-SADS Mania Rating Scale. Results may be analyzed for similarities and differences. Ratings of younger and older populations obtained from a single instrument should be compared. Finally, the impact of comorbidity (e.g., ADHD) on ratings needs to be calculated. That is, does a child with comorbid ADHD and mania score higher on the Y-MRS than one with mania alone, and is that likely to affect the sensitivity to change?

Although Forum participants noted the need to measure impairing features of mania in addition to psychosis, they acknowledged that excessive measures become burdensome, expensive, and possibly dilute the results of the primary outcome measure. It is clear, however, that in contrast to adult mania where psychosis is the principal cause of impairment, in children and adolescents, mania-associated impairment is due principally to irritability and aggression. Assessing both irritability and aggression in youth is particularly important, and participants pointed out the need for further work on measuring these key symptoms. The anchored irritability and aggression items on the Y-MRS similarly were not felt to be sensitive measures. The Overt Aggression Scale-Modified (Coccaro et al., 1991) has been used in several multisite studies of aggression (e.g., McCracken et al., 2002). Although the Overt Aggression Scale-Modified is useful, new or refined aggression scales are needed that measure the frequency, intensity, and duration of rages.

Discussion also considered whether to measure ADHD when it co-occurs. Although the Y-MRS may differentiate large groups of patients with ADHD and mania from patients with ADHD alone, it is unclear how much of the Y-MRS score is accounted for by ADHD. Industry representatives voiced concern about how to interpret a decline in ADHD measures during the course of a clinical trial for mania, especially because patients will be withdrawn from their ADHD medications prior to starting the antimanic treatment. However, conferees felt that ADHD symptom ratings (inattention, hyperactivity, and impulsivity) might usefully be measured at the beginning and endpoints of the experimental antimanic treatment. Alternatively, ADHD symptom ratings might be reserved for euthymic patients, or parent informants could be encouraged to try to rate ADHD as it occurred prior to the onset of mania. Conferees also felt that measures of cognition and academic performance should be taken at baseline and again at a yet-to-be-determined future time (e.g., at the end of the extension phase of a clinical trial). Collaborative research with neuropsychologists is needed to develop appropriate cognitive measures for children and adolescents with acute mania.

Finally, participants agreed that it is necessary to assess a very broad range of family, social, and academic functions among the secondary outcome measures. Pending the development of new instruments, existing general pediatric health outcome measures such as the Child Health Questionnaire (Landgraf et al., 1996) might be useful. The Child Health Questionnaire has been used in studies of ADHD, for example, to look at the impact of the disorder on the family (Michelson et al., 2001).

Multiple sources of information are problematic for weekly ratings of outpatients, just as they are for measures of screening and assessment. Agreement of two out of three respondents increases confidence in actual ratings. A suggestion was made to weight severity on the basis of how many observers note a behavior. If, for example, a parent describes euphoria and the interviewer observes it, the weight should be greater than if only the interviewer had observed euphoria.

Workgroup IV: Design and Ethical Issues

The participation of children and adolescents in a clinical trial adds a greater measure of sensitivity to ethical issues involved in the clinical research. It raises unique protocol design questions, including the use of placebos

and of adjunctive medications used to treat comorbid conditions such as ADHD. Workgroup IV discussed these and other questions and considered the specific needs of subjects and families participating in a clinical trial of a medication to treat bipolar disorder (acute mania).

Consensus Recommendations:

- A placebo arm is needed for Bipolar I disorder studies (mania), with adequate protection for children participating, as noted below.
- Single-therapy (placebo vs. active agent) designs are favored over "add-on" designs, but both are needed.
- The optimal duration for a study of acute mania is at least 3–4 weeks (possibly longer if lithium is involved). Longer studies may put severely ill children on placebo at unacceptably high risk.
- Strategies for recruiting subjects and retaining them in trials include use of inpatient sites and day hospital sites where indicated as well as supports for keeping children and families safe at home if the child is enrolled in an outpatient or day-hospital-based trial. "Family-friendly" policies, such as psychoeducation groups, and therapeutic day school are critically important in clinical trials in outpatient and day hospital settings.
- Studies with children impose special responsibilities on investigators and sponsors, including following federal and local requirements of consent/assent specific to the age of the child participant in the study, identifying the medication used during the treatment trial after the subject has completed the trial in order to provide guidance as to further treatment options (i.e., "breaking the blind"), including an open-label follow-up study after the acute study in successfully treated participants, and requiring publication of negative results.
- A Data Safety Monitoring Board should be instituted for multisite trials involving children and adolescents with acute mania.

Discussion:

The advantages and disadvantages to placebo controls are well known and not unique to children and adolescents (Charney et al., 2002; Laughren, 2001). In spite of the fact that it is felt to be unethical to do placebo-controlled trials in some countries, the group concurred that exposure to placebo is scientifically necessary and ultimately diminishes exposure to potentially ineffective pharmaceutical treatments with serious potential side effects. For essentially all psychiatric disorders, the current scientific standard for an adequate and well-controlled trial is one that is capable of showing a difference between the new drug of interest and a control condition, and this is most efficiently done in a placebo-controlled trial (Laughren, 2001). Even if certain drugs are established as effective in the treatment of mania in pediatric patients, it will still be necessary in trials of new agents to have a placebo arm, because the placebo response rate is unpredictable from trial to trial. If placebo-controlled trials are judged by the community to be unethical once effective treatments are established, the only alternative is the add-on study, in which new drug or placebo is added on to standard treatment. Forum participants expressed enthusiasm for add-on studies, but they generally agreed that the preferred first acute mania study should be a 3- to 4-week, double-blind, monotherapy study with at least 1 year of follow-up in open-label extension in which adjunctive medicines are allowed. The brief duration of the acute trial (3–4 weeks) is designed to protect seriously ill children in the placebo arm and to minimize placebo responding in hospitalized children. The placebo response rate appears to increase with the duration of the trial because of an increased likelihood of spontaneous remission and the response to hospitalization.

The use of placebo controls with outpatients was of concern to all participants, who strongly agreed that "adequate protections" are essential. A Data Safety Monitoring Board[iv] should be required to ensure the safe conduct of the study and to ensure that accurate meaningful data are collected. Interim analyses by the Data Safety Monitoring Board should determine if a study should continue or be terminated because of safety or lack of efficacy concerns. Continuous daily and weekly access to emergency psychiatric evaluation and care, including access to inpatient care, is an essential requirement. Rescue medications during the acute trial pose somewhat of a problem because there are no data that the benzodiazepines, which are used as rescue medications in adult trials, have any efficacy in adolescents, and they are known to disinhibit younger children. No consensus was reached on alternative rescue medications. Conferees agreed that adjunctive medications (e.g., stimulants and antidepressants) need to be discontinued before subjects enter a clinical trial. Although there was broad acknowledgment that this requirement might be a disincentive, if not a deterrent for some potential participants, it did not break the consensus.

In the absence of consensus on rescue medications, participants discussed the use of nonmedication adjunctive treatments and other options that might be made available to patients and families in outpatient and day-hospital-based, placebo-controlled trials. Psychosocial treatments that might serve as adjunctive supports to

[iv]A Data Safety Monitoring Board is an independent group of advisors with varying types of expertise (e.g., pediatrics, psychiatry, lay person/consumer, statistics) convened by the sponsor to review ongoing data on safety and efficacy as well as conduct of the trial.

children and families include psychoeducation groups, social rhythm therapy, personal trainers, nutritional counseling, exercise programs, and regular supportive family counseling. A recommendation was made to include, on pharmaceutical companies' advisory boards, parents of children with the disorder being studied in order to benefit from their insight and advice on issues surrounding study design and recruitment. As part of the ethical conduct of a clinical trial of this type, sponsors should budget for an open-label follow-up of the acute trial, with a specified number of outpatient counseling visits for the patient and family. Psychoeducational testing (e.g., an adaptive function scale like the Vineland Adaptive Behavior Scale; Sparrow et al., 1984) and an IQ test to qualify children for an individualized education plan would be useful as part of the initial assessment and a plan for follow-up care.

Conference participants also discussed how soon maintenance trials should be initiated and what might be reasonable designs for open-label studies. Maintenance studies with adults are generally not conducted until data are available from acute efficacy studies on dosing.

Research Needs

The consensus recommendations summarized in the preceding sections provide guidance to sponsors, regulators, and investigators on the opportunities and limitations involved in adapting clinical trials methodology from studies of acute mania in adults to studies in adolescents and children. It is important to reiterate that because there have been as yet no completed placebo-controlled trials of acute mania in children and adolescents, there remain important gaps in knowledge. These knowledge gaps can help define a research agenda designed to better characterize the different forms of the disorder across younger age groups and identify specific needs for assessment tools and further refinements in inclusion and exclusion criteria.

Improving Diagnostic and Assessment Tools:

As modified and adapted to children, the Y-MRS is currently the best available instrument to measure symptoms of mania in children, but it is far from satisfactory and needs to be improved or replaced. Urgent need exists to translate adult anchors on the Y-MRS into an appropriate child/adolescent version, but the identification of appropriate anchors is contingent on the age of the child and on how the illness/diagnosis is defined. Several researchers currently are revising the Y-MRS while others are attempting to develop improved alternative scales. The opportunity to study other mania measures (e.g., the K-SADS Mania Rating Scale) and health outcomes measures simultaneously would enhance the value of the first generation of placebo-controlled trials in this area. Additional validation studies are needed for diagnostic interviews (e.g., K-SADS).

Characterizing the Population:

Although forum participants agreed broadly on the need to conduct initial studies in children and adolescents with well-defined episodes of acute mania that closely mirror the condition in adults, there was a strong consensus that children who currently carry the diagnosis of BP-NOS are much more common than children who meet DSM-IV criteria for acute manic episode. Participants broadly agreed on the need for separate treatment studies of children/adolescents with a diagnosis of BP-NOS. Criteria for BP-NOS require further deliberation, and more data are needed before treatment studies involving BP-NOS are conducted for purposes of regulatory approval.

Child and adolescent patients participating in clinical trials should be well characterized in terms of family history and course, and investigator-initiated research studies of outcome should be conducted. These can commence with open-label studies, parallel (in time) to the acute Bipolar I mania studies, and generate needed information about phenomenology, laboratory measures, family history, and prognosis. Children and adolescents with BP-NOS could be included in the open-label extensions of the acute bipolar studies, because this phase of the research will not be part of the FDA submission for the treatment of acute mania.

For the longer term, it is critical for the National Institute of Mental Health to support studies of putative biological markers of mania and BP-NOS that can serve to validate the diagnosis and/or indicate treatment response. Neuroimaging and neurophysiological techniques and advances in molecular neurobiology and genetics are promising technologies in this regard.

Addressing Specific Symptom Patterns:

Psychopharmacology practice is increasingly characterized by the use of multiple drugs aimed at specific symptoms within diagnostic categories and/or comorbid disorders. There is growing interest in the possibility of dissecting DSM entities into component symptom complexes and developing treatments for these narrower clinical targets, yet drug development efforts continue at present to focus on monotherapy for DSM-IV-defined Axis I diagnostic entities. Among the types of research needed to move ahead are studies to better define symptoms that occur within specific disorders and to operationalize the assessment of the symptoms. The experimental treatment of specific symptoms should be targeted within the context of discrete disorders. How-

ever, if it can be established that certain symptoms or symptom complexes are nonspecific, in the sense that they occur in identical form in several distinct diagnostic conditions, and studies demonstrate that these symptoms respond to the same treatments, independent of the underlying diagnostic condition in which they occur (e.g., as with pain), the symptoms or symptom pattern could become a primary target in future studies. Research is needed to determine under what conditions studies of specific symptoms (e.g., aggression, mood dysregulation), within well-operationalized conditions (e.g., ADHD, autism/pervasive developmental disorder, conduct disorder, or oppositional defiant disorder), might be appropriate for seeking and developing an indication. In particular, it will be important to secure data on the reliability (and validity) of instruments being used to assess both specific symptoms occurring within the context of discrete diagnostic conditions and also nonspecific symptoms that occur in several diagnostic conditions and can be considered independent of that underlying condition. Research is needed to clarify further these questions before approaching the FDA for a study targeting discrete symptoms seen in BP-NOS.

Summary

With the exception of differences between FDA and CPMP requirements identified by the participant from the European Union, conferees achieved consensus on the major methodological questions that they addressed across the two days of deliberations. The first key point of consensus was the decision to focus initial studies on the treatment of acute manic episodes in children/adolescents aged 10–17 years who may or may not be hospitalized and who may or may not suffer from common comorbid psychiatric disorders such as ADHD, conduct disorder, and substance abuse. Conferees agreed on the requirement that bipolar diagnoses be made by a well-trained child psychiatrist or psychologist, based on information culled from interviews with the child and at least one other source (e.g., parent). Traditional rating scales utilized in adult trials (e.g., Y-MRS) have serious limitations when applied to children and adolescents, but the Y-MRS (supplemented by other rating instruments) was judged to be the best current assessment tool of severity of manic symptoms in children and adolescents. Objective measures and standards for interviewer and rater competency should be established and monitored over the course of the study. Most important, new tools should be developed for assessment, including scales to measure aggression/rage and cognitive function.

Although the participation of children and adolescents in a clinical trial presents important methodological and ethical challenges, placebo-controlled trials are the standard that should be followed. The conferees laid out a set of requirements for the ethical conduct of these studies that should serve as a guidepost to sponsors and investigators. Priority must be given to protecting the health, safety, and well-being of the child and members of the family.

Various research needs were identified. One of the most critical issues, in terms of linking clinical trials to improved clinical practice, involves a better definition of BP-NOS. At present, many of the children and adolescents being treated with mood-stabilizing drugs are being treated for BP-NOS, but the criteria for this disorder are not sufficiently well developed to assess treatment efficacy. One suggestion was to circulate videotapes or vignettes of various types of "not-otherwise-specified" to begin to develop consensus regarding the not-otherwise-specified and mixed-state types

Overall, conferees agreed that studies should proceed on the treatment of acute manic episodes in children and adolescents. This is a vulnerable population needing evidenced-based treatments. Methodologically rigorous large-scale studies will provide vital information regarding the safety and efficacy of drugs with potential mood-stabilizing properties, and treatments that withstand the rigors of scientific scrutiny must be delivered to these patients in hopes of giving them the chance of a brighter future.

Consensus Development Panel Co-Chairs: Gabrielle A. Carlson, MD, Peter S. Jensen, MD, Robert L. Findling, MD, and Roger E. Meyer, MD

Speakers: Marilyn Benoit, MD, Gabrielle A. Carlson, MD, Thomas Laughren, MD, Renee Simar, PhD, David Shaffer, MD, Joseph Calabrese, MD, Eric Youngstrom, PhD, Ellen Leibenluft, MD, and James McCracken, MD

Workgroups:

Workgroup I—Inclusion/Exclusion Criteria: Graham Emslie, MD, Vivek Kusumakar, MD (Co-Chairs); Editha Nottelmann, PhD (Rapporteur); Boris Birmaher, MD, Gwendolyn Boss, Thomas Cummins, MD, Prudence Fisher, PhD, Jean Frazier, MD, Frederick Goodwin, MD, Lydia Lewis, Stephen Murray, Rafael Muniz, MD, Daniel vanKammen, MD, and Jitschak Storosum, MD

Workgroup II—Training and Site Selection: Elizabeth Weller, MD, Janet Wozniak, MD (Co-Chairs); Laurie Flynn (Rapporteur); Albert J. Allen, MD, PhD, Pablo Davanzo, MD, Cheryl Impellizzeri, Mary Kujawa, MD, PhD, Thomas Laughren, MD, Jon McClellan, MD, Theodore Petti, MD, David Shaffer, MD, Renee Simar, PhD, and Alan Unis, MD

Workgroup III—Outcome Measures: Michael Strober, PhD, Gary Sachs, MD (Co-Chairs); Robert Kowatch, MD (Rapporteur); David Axelson, MD, Melanie Bruno, PhD, MBA, Kiki Chang, MD, Joseph D'Souza, PhD, Reuven Ferziger, MD, Wendy Granberry, MT, RPH, Darcy Gruttadaro, JD, Gahan Pandina, PhD, and Eric Youngstrom, PhD

Workgroup IV—Design and Ethical Issues: Daniel Pine, MD, Melissa DelBello, MD (Co-Chairs); Martha Hellander (Rapporteur); Robinder Bhangoo, MD, Joseph Calabrese, MD, Srihari Gopal, MD, Vivian Kafantaris, MD, Ludmila Kryzhanovskaya, MD, PhD, Robert Leadbetter, MD, James McCracken, MD, Brian Miller, PharmD, BCPP, Kathryn Tracy, MD, PhD, Benedetto Vitiello, MD, and Lamia Wassef, MD

Conference Sponsors: American Academy of Child and Adolescent Psychiatry (Virginia Q. Anthony, Executive Director) and Best Practice (Roger E. Meyer, MD, CEO)

Conference Underwriters: Abbott Laboratories, Bristol-Myers Squibb Company, GlaxoSmithKline, INC Research, Janssen Pharmaceutica, Johnson & Johnson Pharmaceutical Research and Development, Eli Lilly and Company, Pfizer Inc., Novartis Pharmaceuticals Corporation, Solvay Pharmaceuticals, Inc.

Support Staff: Patricia Brown (American Academy of Child and Adolescent Psychiatry) and Arthur Gold (Best Practice)

Writer: Paul Sirovatka, MS

Author and Contributor Affiliations

Albert Einstein College of Medicine, Bronx, New York (Dr. Kafantaris); Case Western Reserve University, Cleveland, Ohio (Drs. Calabrese, Findling, and Youngstrom); Columbia University, New York, New York (Drs. Jensen, Fisher, and Shaffer and Ms. Flynn); Dalhousie University, Halifax, Nova Scotia, Canada (Dr. Kusumakar); Harvard University, Boston, Massachusetts (Drs. Frazier, Sachs, and Wozniak); Stanford University, Stanford, California (Dr. Chang); Stony Brook University, Stony Brook, New York (Dr. Carlson); University of California-Los Angeles (Drs. Davanzo, McCracken, and Strober); Indiana University School of Medicine, Indianapolis, Indiana (Dr. Petti); University of Texas Southwestern Medical Center, Dallas, Texas (Dr. Emslie); George Washington University, Washington, DC (Dr. Goodwin); University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania (Drs. Axelson and Birmaher); Food and Drug Administration, Rockville, Maryland (Dr. Laughren[1]); National Institute of Mental Health, Bethesda, Maryland (Drs. Liebenluft, Nottelmann, Pine, and Vitiello); Best Practice, Bethesda, Maryland (Drs. Meyer and Goodwin); Child and Adolescent Bipolar Foundation, Wilmette, Illinois (Ms. Hellander); Depression and Bipolar Support Alliance, Chicago, Illinois (formerly National DMDA) (Lydia Lewis); INC Research, Austin, Texas (Dr. Simar); University of Cincinnati, Cincinnati, Ohio (Drs. DelBello and Kowatch); University of Pennsylvania-Children's Hospital of Philadelphia (Dr. Weller); National Alliance for the Mentally Ill, Arlington, Virginia (Ms. Gruttadaro); Abbott Laboratories, Abbott Park, Illinois (Thomas Cummins, MD, and Kathym Tracy, MD, PhD); Bristol-Myers Squibb, Plainsboro, New Jersey (Cheryl Impellizzeri and Mary Kujawa, MD, PhD); Eli Lilly and Company, Indianapolis, Indiana (Albert J. Allen, MD, PhD, Melanie A. Bruno, PhD, MBA, and Ludmilla Kryzhanovskaya, MD, PhD); GlaxoSmithKline, Research Triangle Park, North Carolina (Wendy Granberry, MT, RPH, and Robert Leadbetter, MD); Janssen Pharmaceutical Products, Titusville, New Jersey (Reuven Ferziger, MD, and Gahan Pandina, PhD); Johnson & Johnson Pharmaceutical Research and Development, Raritan, New Jersey (Sri Gopal, MD, Allan Unis, MD, and Daniel vanKammen, MD); Novartis Pharmaceuticals Corporation, East Hanover, New Jersey (Joseph D'Souza, PhD, and Rafael Muniz, MD); Pfizer, Inc., New York, New York (Stephen Murray); Solvay Pharmaceuticals, Inc., Marietta, Georgia (Gwendolyn Boss and Brian Miller, PharmD, BCPP); Medicines Evaluation Board (Jitschak Storosum, MD); and Clinical Trials and Special Access Programme, Ontario, Canada (Lamia Wassef, MD).

Drs. Storosum and Wassef attended as listeners articulating European and Canadian regulatory experience. Dr. Storosum is with the Medicines Evaluation Board of the Netherlands, Den Haag. However, Dr. Storosum's contribution to this article was made in his private capacity. No official support or endorsement by the Medicines Evaluation Board or CPMP is intended or should be inferred. Dr. Wassef is with the Health Canada, Clinical Trials and Special Access Programme.

References

Altman EG, Hedeker DR, Janicak PG, Peterson JL, Davis JM (1994): The Clinician-Administered Rating Scale for Mania (CARS-M): Development, reliability, and validity. *Biol Psychiatry* 36(2):124–134.

American Psychiatric Association (1994): *Diagnostic and Statistical Manual of Mental Disorders,* 4th ed. Washington, DC, American Psychiatric Association.

American Psychiatric Association Task Force for the Handbook of Psychiatric Measures. (2000): *Handbook of Psychiatric Measures.* Washington, DC, American Psychiatric Association.

Asberg M, Montgomery S, Perris C, Schalling D, Sedvall G (1978): A comprehensive psychopathological rating scale. *Acta Psychiatr Scand* (Suppl) 271:5–27.

Bech P, Bolwig TG, Kramp P, Rafaelsen OJ (1979): The Bech-Rafaelsen Mania Scale and the Hamilton Depression Scale. *Acta Psychiatr Scand* 59(4):420–430.

Beigel A, Murphy DL (1971): Assessing clinical characteristics of the manic state. *Am J Psychiatry* 128(6):688–694, 1971.

Biederman J, Klein RG, Pine DS, Klein DF (1998): Resolved: Mania is mistaken for ADHD in prepubertal children. *J Am Acad Child Adolesc Psychiatry* 37(10):1091–1096; discussion 1096–1099.

Blackburn IM, Loudon JB, Ashworth CM (1977): A new scale for measuring mania. *Psychol Med* 7(3):453–458.

Bowden CL, Brugger AM, Swann AC, Calabrese JR, Janicak PG, Petty F, Dilsaver SC, Davis JM, Rush AJ, Small JG, Garza-Trevino ES, Risch SC, Goodnick PJ, Morris DD (1994): Efficacy of divalproex vs. lithium and placebo in the treatment of mania. The Depakote Mania Study Group. *JAMA* 271(12):918–924.

Brierley CE, Szabadi E, Rix KJ, Bradshaw CM (1988): The Manchester Nurse Rating Scales for the daily simultaneous assessment of depressive and manic ward behaviours. *J Affect Disord* 15(1):45–54.

Campbell M, Small AM, Green WH, Jennings SJ, Perry R, Bennett WG, Anderson L (1984): Behavioral efficacy of haloperidol and lithium carbonate: A comparison in hospitalized aggressive children with conduct disorder. *Arch Gen Psychiatry* 41:650–656.

Carlson GA (1998): Mania and ADHD: Comorbidity or confusion? [review]. *J Affect Disord* 51(2):177–187.

Carlson GA (1999): Juvenile mania versus ADHD. *J Am Acad Child Adolesc Psychiatry* 38(4):353–354.

Chambers WJ, Puig-Antich J, Hirsch M, Paez P, Ambrosini PJ, Tabrizi MA, Davies M (1985): The assessment of affective disorders in children and adolescents by semi-structured interview. *Arch Gen Psychiatry* 42:696–702.

Charney DS, Nemeroff CB, Lewis L, Laden SK, Gorman JM, Laska EM, Borenstein M, Bowden CL, Caplan A, Emslie GJ, Evans DL, Geller B, Grabowski LE, Herson J, Kalin NH, Keck PE Jr, Kirsch I, Krishnan KR, Kupfer DJ, Makuch RW, Miller FG, Pardes H, Post R, Reynolds MM, Roberts L, Rosenbaum JF, Rosenstein DL, Rubinow DR, Rush AJ, Ryan ND, Sachs GS, Schatzberg AF, Solomon S (2002): Consensus Development Panel: National Depressive and Manic-Depressive Association consensus statement on the use of placebo in clinical trials of mood disorders. *Arch Gen Psychiatry* 59(3):262–270.

Coccaro EF, Harvey PD, Kupsaw-Lawerence E, Herbert JL, Bernstein DP (1991): Development of neuropharmacologically based behavioral assessments of impulsive aggressive behavior. *J Neuropsychiatry Clin Neurosci* 3:S44–S51.

Committee for Proprietary Medicinal Products (CPMP) (2000): *Note for Guidance on Clinical Investigation of Medicinal Products in the Paediatric Population* (CPMP/ICH/2711/99-adopted in July 2000). [http:// www.emea.eu.int/htms/human/ich/efficacy/ichfin.htm].

Committee for Proprietary Medicinal Products (CPMP) (2001): *Note for Guidance on Clinical Investigation of Medicinal Products for the Treatment and Prevention of Bipolar Disorder* (CPMP/EWP/567/98-adopted in April 2001). [http:// www.emea.eu.int/pdfs/human/ewp/056798en.pdf].

Emslie G, Rush JA, Weinberg WA, Kowatch RA, Hughes CW, Carmody T, Rintelmann J (1997): A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Arch Gen Psychiatry* 54: 1031–1037.

Endicott V, Spitzer RL (1978): A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia. *Arch Gen Psychiatry* 35:837–848.

First MB, Gibbon M, Spitzer-Williams N, Spitzer RL, Williams JBW (1995): *Structured Clinical Interview for DSM-IV (SCID-I) (Users Guide and Interview) Research Version*. New York, Biometrics Research Department, New York Psychiatric Institute.

Guy W (1976): ECDEU *Assessment Manual for Psychopharmacology—Revised* (DHEW Publ No ADM 76-388). Rockville, MD, U.S. Department of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, NIMH Psychopharmacology Research Branch, Division of Extramural Research Programs, pp. 218–222.

Hamilton M (1960): A rating scale for depression. *J Neurol Neurosurg Psychiatry* 23:56–62.

Heinig SJ, Quon ASW, Meyer RE, Korn D (1999): The changing landscape for clinical research. *Acad Med* 74:725–745.

Herjanic B, Reich W (1977): Differentiating psychiatrically disturbed children on the basis of a structured interview. *J Abnorm Child Psychol* 5:127–134.

Herjanic B, Reich W (1982): Development of a structured psychiatric interview for children: Agreement between child and parent on individual symptoms. *J Abnorm Child Psychol* 10:307–324.

Hughes CW, Rintelmann J, Emslie GJ, Lopez M, MacCabe N (2001): A revised anchored version of the BPRS-C for childhood psychiatric disorders. *J Child Adolesc Psychopharmacol* 11(1):77–93.

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N (1997): Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 36:980–989.

Kay SR, Fiszbein A, Opler LA (1987): The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr Bull* 13:261–276.

Kowatch RA, Suppes T, Carmody TJ, Bucci JP, Hume JH, Kromelis M, Emslie GJ, Weinberg WA, Rush AJ (2000): Effect size of lithium, divalproex sodium, and carbamazepine in children and adolescents with bipolar disorder. *J Am Acad Child Adolesc Psychiatry* 39(6):713–720.

Landgraf JM, Abetz L, Ware JE (1996): *The CHG User's Manual*, 1st ed. Boston, Health Institute, New England Medical Center.

Laughren TP (2001): The scientific and ethical basis for placebo-controlled trials in depression and schizophrenia: An FDA perspective. *Eur Psychiatry* 16(7): 418–423.

McCracken JT, McGough J, Shah B, Cronin P, Hong D, Aman MG, Arnold LE, Lindsay R, Nash P, Hollway J, McDougle CJ, Posey D, Swiezy N, Kohn A, Scahill L, Martin A, Koenig K, Volkmar F, Carroll D, Lancor A, Tierney E, Ghuman J, Gonzalez NM, Grados M, Vitiello B, Ritz L, Davies M, Robinson J, McMahon D (2002): Research Units on Pediatric Psychopharmacology Autism Network. Risperidone in children with autism and serious behavioral problems. *N Engl J Med* 347(5):314–321.

Michelson D, Faries E, Wernicke J, Kelsey D, Kendrick K, Sallee R, Spencer T, and the Atomoxetine Study Group (2001): Atomoxetine in the treatment of children and ado-

lescents with attention-deficit/hyperactivity disorder: A randomized, placebo-controlled, dose-response study. *Pediatrics* 108:83–91.

Montgomery SA, Asberg M (1979): A new depression scale designed to be sensitive to change. *Br J Psychiatry* 134: 382–389.

Overall JE, Gorham DR (1988): The Brief Psychiatric Rating Scale (BPRS): Recent developments in ascertainment and scaling. *Psychopharmacol Bull* 24:97–99.

Poznanski EO, Cook SC, Carroll BJ, Corzo H (1983): Use of the Children's Depression Rating Scale in an inpatient psychiatric population. *J Clin Psychiatry* 44:200–203.

Poznanski EO, Freeman LN, Mokros HB (1985): Children's Depression Rating Scale-Revised (September, 1984). *Psychopharmacol Bull* 21:979–989.

Shaffer D, Fisher P, Dulcan MK, Davies M, Piacentini J, Schwab-Stone ME, Lahey BB, Bourdon K, Jensen PS, Bird HR, Canino G, Regier DA (1996): The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): Description, acceptability, prevalence rates, and performance in the MECA Study. Methods for the Epidemiology of Child and Adolescent Mental Disorders Study. *J Am Acad Child Adolesc Psychiatry* 35:865–877.

Sparrow W, Baila DA, Cicchetti DV (1984): *Vineland Adaptive Behavior Scales*. Circle Pines, Minnesota, American Guidance Service, Inc.

Young RC, Biggs JT, Ziegler VE, Meyer DA (1978): A rating scale for mania: Reliability, validity, and sensitivity. *Br J Psychiatry* 133:429–435.

Zisson S (2001, December): "GSK Forges Ties with Academia." CenterWatch 8(12), pp. 1, 8–12.

Address correspondence to:
Gabrielle A. Carlson, MD
Department of Psychiatry
Stony Brook University-Putnam Hall
Stony Brook, NY 11794-8790

E-mail: Gabrielle.Carlson@StonyBrook.edu

Received for publication December 2, 2002; accepted after revision December 16, 2002.

Published online: January 24, 2003.