Assessing Cognitive Function in Bipolar Disorder: Challenges and Recommendations for Clinical Trial Design

Katherine E. Burdick, PhD; Terence A. Ketter, MD; Joseph F. Goldberg, MD; and Joseph R. Calabrese, MD

ABSTRACT

Objective: Neurocognitive impairment in schizophrenia has been recognized for more than a century. In contrast, only recently have significant neurocognitive deficits been recognized in bipolar disorder. Converging data suggest the importance of cognitive problems in relation to quality of life in bipolar disorder, highlighting the need for treatment and prevention efforts targeting cognition in bipolar patients. Future treatment trials targeting cognitive deficits will be met with methodological challenges due to the inherent complexity and heterogeneity of the disorder, including significant diagnostic comorbidities, the episodic nature of the illness, frequent use of polypharmacy, cognitive heterogeneity, and a lack of consensus regarding measurement of cognition and outcome in bipolar patients. Guidelines for use in designing future trials are needed.

Participants: The members of the consensus panel (each of the bylined authors) were selected based upon their expertise in bipolar disorder. Dr Burdick is a neuropsychologist who has studied cognition in this illness for 15 years; Drs Ketter, Calabrese, and Goldberg each bring considerable expertise in the treatment of bipolar disorder, both within and outside of controlled clinical trials. This consensus statement was derived from work together at scientific meetings (eg, symposium presentation at the 2014 Annual Meeting of the American Society of Clinical Psychopharmacology, among others) and ongoing discussions by conference call. With the exception of the public presentations on this topic, these meetings were closed to outside participants.

Evidence: A literature review was undertaken by the authors to identify illness-specific challenges relevant to the design and conduct of treatment trials targeting neurocognition in bipolar disorder. Expert opinion from each of the authors guided the consensus recommendations.

Consensus Process: Consensus recommendations, reached by unanimous opinion of the authors, are provided here as a preliminary guide for future trial design. Recommendations comprise exclusion of certain syndromal-level comorbid diagnoses and current affective instability, restrictions on numbers and types of medications, and use of prescreening assessment to ensure enrollment of subjects with adequate objective evidence of baseline cognitive impairment.

Conclusions: Clinical trials to address cognitive deficits in bipolar disorder face distinctive design challenges. As such trials move from proof-of-concept to confirmation of clinical efficacy, it will be important to incorporate distinctive design modifications to adequately address these challenges and increase the likelihood of demonstrating cognitive remediation effects. The field is now primed to address these challenges and increase the likelihood of demonstrating cognitive remediation effects. The field is now primed to address these challenges and increase the likelihood of demonstrating cognitive remediation effects.

METHOD

This article briefly reviews studies of cognitive dysfunction in bipolar disorder, summarizes the challenges inherent in the design and conduct of treatment trials targeting neurocognition in patients with bipolar disorder, and provides recommendations for addressing these challenges. The review of cognition in bipolar disorder is not comprehensive, as the focus of this article was on consensus development; however, for a recent full review of this area incorporating associated neurobiological changes seen in bipolar disorder, see Lim et al.8

The consensus process utilized for the development of this work entailed the joint effort of all 4 bylined authors, each with expertise in bipolar disorder, clinical trials, and neurocognitive functioning. A series of meetings, primarily by conference call, resulted in the recommendations put forth in this article. The primary limitation to this process is the fact that these were...
There has been a strong focus on cognitive deficits in schizophrenia over the past century, with concerted efforts put forth to begin targeting them directly for treatment over the past few decades. In contrast, parallel efforts in bipolar disorder are lagging, despite convergent data indicating direct and deleterious effects on functional outcome and quality of life.

This article is aimed at researchers designing clinical trials for cognition in bipolar patients; however, lessons learned from such controlled trials can translate directly to patient care, with the ultimate goal of developing new agents to combat this disabling aspect of the illness. Moreover, a better understanding of the illness features that contribute to the neurocognitive problems in this disorder will allow for more meaningful psychoeducation for patients who hope to limit and even prevent cognitive dysfunction from occurring.

Cognitive Dysfunction in Bipolar Disorder

Recent years have seen exponential growth in studies of cognitive dysfunction in bipolar disorder (Figure 1), consistent with the current NIMH focus on implementing a dimensional approach to neuropsychiatric illnesses with the Research Domain Criteria (RDoC) initiative (http://www.nimh.nih.gov/research-priorities/rdoc/index.shtml). Neurocognitive functioning represents one of the key RDoC constructs that crosses DSM-5 boundaries and may substantially enhance our understanding of the pathophysiology of diverse brain-based illnesses, including bipolar disorder.

Neurocognitive deficits have long been acknowledged as a core feature in schizophrenia. Emil Kraepelin differentiated dementia praecox (schizophrenia) from manic-depressive psychosis (bipolar disorder) in the early 20th century, believing that patients with bipolar disorder exhibited affective and cognitive symptom-free, euthymic intervals between mood episodes. More recent data indicate that a large proportion of patients with bipolar disorder experience only partial recovery from affective and cognitive symptoms between episodes. Longitudinal studies have demonstrated that many patients experience ongoing subsyndromal affective symptoms and that neurocognitive deficits are prominent during euthymia. In the cognitive sphere, persistent, trait-like deficits have been noted in attention, verbal learning, and executive function, with mean performance falling approximately 1 standard deviation below that of healthy controls.

Studies suggest that cognitive impairments in bipolar disorder qualitatively overlap with, but are less severe than, those in schizophrenia. One notable difference is that impairment in global cognitive performance (eg, IQ) is characteristic of schizophrenia but generally absent in bipolar disorder. This may reflect the differences noted in premorbid IQ. Whereas future bipolar patients tend to show normal-range IQ before illness onset, many schizophrenia patients have IQ impairment well before the emergence of psychotic symptoms.

However, it is now widely recognized that deficits in more specific neurocognitive domains occur early on in bipolar disorder and escalate with an increasing number of major affective episodes. Cognitive impairment occurs in both bipolar I and II disorders, although some, but not all, studies indicate more impairment in bipolar I disorder. In addition, meta-analytic data suggest that bipolar patients with a history of psychosis (which is more common in bipolar I than bipolar II disorder) perform worse than bipolar patients who have never experienced psychosis, even during affective remission.

An important finding, as shown repeatedly in schizophrenia and bipolar disorder, is that persistent neurocognitive deficits in bipolar disorder contribute significantly to functional disability that prevents patients from achieving complete functional recovery. Although bipolar disorder treatment studies have previously focused on affective symptoms, it is important to consider the substantial disability associated with impaired cognition. Even though a recent focus has been placed on the importance of carrying out large-scale clinical trials to specifically target cognitive impairment in patients with schizophrenia, an analogous approach in bipolar disorder remains to be realized.

Figure 1. Exponential Increase in Publications Related to Cognition in Bipolar Disorder Over Time*
Assessing Cognitive Function in Bipolar Disorder

Challenges and Recommendations for Trial Design in Bipolar Disorder

The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) initiative, a collaborative effort among academia, the US Food and Drug Administration (FDA), and the pharmaceutical industry, led to the development of guidelines for study design and a consensus cognitive battery to measure cognitive outcomes in clinical trials in patients with schizophrenia. Several recommendations were made for optimizing outcomes in clinical trials in patients with schizophrenia. These recommendations entail substantive differences from those in schizophrenia patients and necessary efforts such as the MATRICS initiative have not yet been realized in bipolar disorder. Challenges in bipolar disorder research include substantial heterogeneity in illness subtype and course (e.g., type I vs type II, rapid cycling), fluctuating mood states over time, diverse common comorbid conditions (e.g., substance abuse, anxiety disorders), and frequent need for polypharmacy. Cognitive heterogeneity (i.e., patients ranging from unimpaired to very impaired) is also a concern in bipolar trials but is less of an issue in schizophrenia. Table 1 presents some of these challenges and suggests preliminary recommendations, which are discussed in detail below.

Subject Characteristics

The updated MATRICS guidelines recommend less restrictive exclusion criteria than originally proposed to facilitate patient recruitment while preserving study validity. Such an approach will most likely also apply when enrolling patients with bipolar disorder to optimize signal detection while permitting feasible recruitment and generalizable results. Several bipolar-specific issues should be carefully considered when designing cognitive trials.

### Table 1. Summary of Challenges and Recommendations for Cognitive Trial Design in Bipolar Disorder

<table>
<thead>
<tr>
<th>Challenge/Confound</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagnosis issue</strong></td>
<td>Include both subtypes, with cognitive impairment threshold</td>
</tr>
<tr>
<td>Bipolar subtype (bipolar I/II disorder)</td>
<td>Allow history of substance use disorder; exclude very recent history; test for illicit substances at time of assessment to rule out intoxication</td>
</tr>
<tr>
<td>Substance use disorder comorbidity</td>
<td>Allow anxiety disorder diagnosis; evaluate current state and rule out subjects with current syndromal anxiety symptoms</td>
</tr>
<tr>
<td>Anxiety disorder comorbidity</td>
<td>Allow subjects who received ADHD diagnosis in adulthood (after bipolar onset), but exclude subjects who received diagnosis and treatment in childhood</td>
</tr>
<tr>
<td>ADHD comorbidity</td>
<td>Institute maximum threshold for both depression and mania symptoms; exclude acute mania/depression; subsyndromal symptoms allowable</td>
</tr>
<tr>
<td>Clinical issue</td>
<td>Adjunctive design recommended, add-on to current mood-stabilizing regimen</td>
</tr>
<tr>
<td>Affective symptoms (state)</td>
<td>Exclude subjects taking high-dose anticholinergics, topiramate, clozapine, tricyclic antidepressants, benzodiazepines within 6 hours of testing, those who had ECT within past 6 months</td>
</tr>
<tr>
<td>Concomitant medications</td>
<td>Restrict inclusion to subjects taking no more than 3 psychotropic medications; dosing limitations might also be useful</td>
</tr>
<tr>
<td>Specific medication restrictions</td>
<td>Include a prescreening battery to test for objective evidence of impairment; threshold recommended at –1 SD below average</td>
</tr>
<tr>
<td>No. of medications</td>
<td>Should be sensitive enough to detect meaningful change, practical in its implementation, limited in practice effects; should be objective and not based on patient self-report</td>
</tr>
<tr>
<td>Presence/severity of cognitive deficits</td>
<td>The MCCB is a reasonable core battery; additional measures may be required to enhance sensitivity and to tap into the affective processing domain</td>
</tr>
<tr>
<td>Neurocognitive battery</td>
<td>Should be sensitive enough to capture partial disability; should be based on objective information and an informant whenever possible</td>
</tr>
<tr>
<td>Functional outcome measures</td>
<td>The MCCB is a reasonable core battery; additional measures may be required to enhance sensitivity and to tap into the affective processing domain</td>
</tr>
</tbody>
</table>

**Abbreviations:** ADHD = attention-deficit/hyperactivity disorder, ECT = electroconvulsive therapy, MATRICS = Measurement and Treatment Research to Improve Cognition in Schizophrenia, MCCB = MATRICS Consensus Cognitive Battery.
driving under the influence, or more than 2 admissions to therapeutic medical detoxification and/or residential treatment programs spanning at least a decade) should probably be excluded. The current use of certain illicit substances not meeting DSM criteria for substance abuse or dependence should also be considered exclusionary, as is done in pharmacotherapy affective efficacy studies in bipolar disorder in which marijuana is allowable but cocaine and other drugs are excluded. In this case, a toxicology screen on testing days could ensure that acute intoxication is not affecting the primary results. However, the sustained presence of markers of cannabis use in toxicology screens may make the latter challenging. Regardless of the approach used to reduce the confounding effects of substance misuse, standardized measures should be incorporated into trials to ensure adequate measurement of substance use (even within normal range), inclusive of legal substances such as alcohol and tobacco. Anxiety disorders are also highly comorbid with bipolar disorder and could have significant effects on cognitive performance; therefore, the exclusion of patients with current syndromal anxiety disorders should also be considered in cognitive trials.

Finally, a diagnosis of attention-deficit/hyperactivity disorder (ADHD) is common in younger individuals with bipolar disorder. However, some bipolar patients are first diagnosed with ADHD in adulthood after the onset of bipolar disorder, making it difficult to differentiate between cognitive symptoms associated with bipolar illness and ADHD. For the purposes of cognitive trials, it might be useful to exclude bipolar patients who received a diagnosis of ADHD during childhood and who were treated for this condition before the onset of bipolar disorder, rather than exclude all subjects who have ever received a diagnosis of ADHD. This strategy would reduce the likelihood that individuals with true ADHD diagnoses are included in cognitive trials designed to target cognitive problems associated more specifically with bipolar disorder.

Mood State

The episodic nature of bipolar illness requires careful consideration of baseline symptom severity, particularly in light of data suggesting a significant influence of affective symptoms on cognitive performance. If patients with current affective symptoms are included in a cognitive trial and affective improvement is reported, it may be difficult to rule out pseudospecificity. Restricting inclusion to patients who are affectively stable at the time of treatment initiation is one way to substantially reduce confounding due to affective symptoms. One can choose to use strict euthymia criteria or a slightly relaxed stability definition, depending upon practicality and the specific agent being studied. If the medication being tested has possible antidepressant activity or a high risk of inducing mania, it might be preferable to use the more stringent euthymia definition to avoid effects of mood-state changes. However, given the pervasiveness of subthreshold affective symptoms, it may be necessary to permit a modest degree of such symptoms in the interest of enrollment feasibility. The degree to which subthreshold depressive symptoms, compared to mood elevation symptoms, might affect cognition remains to be established; however, early data suggest an interactive effect of subthreshold depression and verbal memory deficits on everyday functioning in patients with bipolar disorder.

Current affective symptom severity should be assessed at screening and baseline and intermittently (especially on cognitive testing days) throughout the trial by using standardized mood ratings, such as the Hamilton Depression Rating Scale or the Montgomery-Asberg Depression Rating Scale and the Young Mania Rating Scale. These ratings can serve to confirm euthymic or stable state at study entry and to track any changes in affective symptoms over the course of the study. In general, it is recommended to enroll subjects who are in a relative state of remission; however, it is important to acknowledge that many patients with bipolar disorder have persistent subsyndromal affective symptoms, particularly low-grade depression, during otherwise “remitted” periods. Thus, in an effort to optimize both feasibility and generalizability, while still reducing the degree to which affective symptoms may confound cognitive analyses, cutoffs for entry may be less restrictive than those defining pure euthymia. The inclusion of cases with subsyndromal symptoms can be dealt with post hoc either by stratifying analyses based upon purely euthymic versus subsyndromal status or by including mood ratings as covariates in statistical models. More sophisticated regression or pathway-based models could eventually contribute important information on the interactions between cognition, mood, and treatment outcome in patients with bipolar illness.

In an effort to limit the confounding effect of mood cycling, the shortest possible trial duration should be considered, while still acknowledging the importance of assessing the relative persistence of any cognitive benefit that might be found. Duration will need to be adjusted based upon the specific agent of interest, its mechanism of action, the presumed onset of efficacy, and the recommended dosing strategy. Because mood-state switches or symptom exacerbations may occur during the trial, a clear threshold or specific criteria should be used for subject discontinuation to limit the degree to which mood-state changes introduce “noise” when attempting to detect a significant cognitive signal. In patients who discontinue pharmacotherapy interventions due to adverse effects (rather than due to inefficacy), off-on-off designs could be informative, although the possibility of lingering postdiscontinuation cognitive effects needs to be addressed.

Another suggestion for trial design when considering mood-state confounding, as well as safety-related concerns, is the required use of a concomitant mood-stabilizing medication. Although monotherapy studies might be considered optimal when testing new agents to control affective symptoms in bipolar disorder, the targeting of neurocognitive function is best done when the patient is already affectively stable, which typically requires pharmacotherapy. In addition, the inclusion
of unmedicated bipolar patients is likely to significantly reduce the generalizability of the study given the widespread use of medications in this disorder. As such, the potential confounding effects of concomitant medications needs to be considered in trial design.

Concomitant Medications
The MATRICS committee recommends add-on study designs in schizophrenia because monotherapy may be neither ethically acceptable nor representative of actual clinical practice. Although this approach might also be desirable in bipolar disorder, it may confound cognitive outcomes in bipolar patients, who commonly take multiple psychotropic medications.

Definite data are lacking with regard to the impact on cognition of commonly used medications for bipolar disorder; however, it is clear that at least some treatments that stabilize mood can negatively affect cognitive performance. For example, lithium may be associated with distinct, although subtle, negative effects in several cognitive domains, including verbal learning, short-term memory, creativity, and psychomotor performance; however, when excluding situations in which there might be some degree of lithium intoxication, several longitudinal studies have not reported such deleterious effects.

Among the anticonvulsant mood stabilizers, valproate is associated with mild impairments in attention, short- and long-term memory, and motor speed; carbamazepine use is reported to result in mild memory impairments. Lamotrigine appears to be cognitively benign, whereas topiramate, which is not indicated for bipolar disorder but may be prescribed for prevention of migraine headaches or weight control, is associated with extensive adverse cognitive effects.

Atypical antipsychotic agents vary in their propensity for sedation (diminished cognitive responsiveness) and somnolence (subjective sleepiness), with the greatest incidence noted for clozapine, olanzapine, and quetiapine. Less-sedating antipsychotic agents, such as aripiprazole and ziprasidone, can yield akathisia, anxiety, and agitation, which could also adversely affect cognition. Lurasidone, a less-sedating atypical antipsychotic recently approved for bipolar I depression, although also associated with akathisia, may have a positive effect on cognition, putatively due to 5-HT and 5-HT receptor binding within the hippocampus.

A recent review evaluated the effects of antidepressants on neurocognitive function and concluded that the sedative and negative cognitive effects of tricyclic antidepressants are well established. Whereas selective serotonin reuptake inhibitors generally have a neutral effect on cognitive performance, sertraline may be somewhat better and paroxetine somewhat worse than other drugs in this class. Bupropion, reboxetine, and selective norepinephrine reuptake inhibitors may have more beneficial effects than other antidepressants. It should be noted that newer agents with novel mechanisms (eg, vortioxetine) may show some promise in improving cognition in acutely depressed patients with major depressive disorder.

Although limited data are available on any specific drug class with regard to the adverse effects of pharmacology in patients with bipolar disorder, there are no studies directly addressing polypharmacy, which is very prevalent in bipolar patients, or the cumulative effects of multiple psychotropic agents on cognitive performance. Furthermore, there have been no controlled cognitive enhancement trials in bipolar disorder that have been large enough to assess the effects of concomitant medications or to stratify randomization to control for these effects. In our recently completed cognitive enhancement study of pramipexole, we were able to determine the effects of concomitant medications only superficially, by dividing groups based on the presence or absence of specific classes of medication (eg, antipsychotics, anticonvulsants, lithium), and we were unable to account for polypharmacy in any meaningful way.

When designing future trials, restrictions with regard to concomitant medications might be considered somewhat arbitrary given the lack of clear evidence of deleterious cognitive effects for most bipolar pharmacotherapies; however, setting some limitations would be prudent. Given the existing evidence, medications such as topiramate, tricyclic antidepressants, clozapine, and high doses of anticholinergics should be disallowed. Moreover, subjects who have received electroconvulsive therapy during the past 6 months should be excluded from cognitive trials. Following MATRICS recommendations, and in an effort to reduce treatment heterogeneity, subjects taking more than 3 psychotropic medications should also be excluded. These limits might be difficult to implement given the very high rates of polypharmacy in bipolar patients. In addition, neuroimaging studies in bipolar patients have adopted an approach to quantify concomitant medication load based upon dosage and relative burden of each medication class on brain function, which could be applied in clinical trial analyses as well. Sedating medications such as benzodiazepines should be restricted within 6 hours of testing, and other restrictions should be considered depending upon the agent being investigated and specific drug-drug interactions. Finally, although little is known about dose-related adverse cognitive effects, it may be prudent to limit participation, where applicable, to subjects who are taking recommended doses of standard mood stabilizers, subjects whose doses and serum concentrations are in the lower half of the therapeutic range, or both groups. Each of these considerations should be carefully weighed while balancing recruitment feasibility and outcome optimization.

Cognitive Heterogeneity in Bipolar Disorder
Another unique methodological issue concerns the considerable cognitive heterogeneity that is seen in bipolar disorder; in fact, a substantial proportion of patients (perhaps as high as 40%) display neuropsychological performance within the normal range. Any study focusing on neurocognitive dysfunction must ensure that a sufficient level of deficit is present at baseline to be adequately targeted.
In cognitive trials of schizophrenia, it has been unnecessary to impose a specific inclusion criterion related to a minimal level of cognitive impairment at baseline; however, there is a potential need to do so in patients with bipolar disorder. Indeed, in our controlled cognitive trial of pramipexole in bipolar patients, preliminary findings indicated that lower baseline cognitive performance was correlated with greater cognitive improvement in the patients who were taking pramipexole. These findings were only present in the active medication group, not in the placebo group, suggesting that this did not represent regression to the mean but rather an actual treatment effect. Moreover, the potential for ceiling effects in some patients with bipolar illness is of real concern when optimizing study design; therefore, it may be of particular importance in bipolar disorder trials to prescreen subjects for objective evidence of baseline cognitive dysfunction and to exclude those patients with minimal or no impairment.

Several potential approaches could be implemented for prescreening patients for sufficient cognitive dysfunction. One approach would be to use patient self-report, or subjective measures, of cognitive impairment; however, this is problematic in patients with bipolar disorder because it has been shown to be only very weakly correlated with objective performance on neurocognitive tasks. In fact, cognitive complaints among bipolar patients tend to be more highly associated with mood symptoms—the extent to which this occurs with subthreshold affective symptoms needs to be established. Objective evidence of cognitive deficit, employing a prescreening battery different from the battery used as the primary outcome measure in a trial, may be more practical. Global screening measures commonly used in other clinical samples (eg, the Mini-Mental State Examination [MMSE]) are most likely not applicable in this context due to low sensitivity. Rather, prescreening batteries should include specific measures that are most commonly impaired in patients with bipolar disorder (eg, verbal memory, attention). The threshold used as a cutoff can also be somewhat arbitrary; however, impairment of at least 0.5 SD below normative performance is typically considered minimally relevant from a clinical standpoint. To permit adequate room for improvement in performance, it may be preferable to require evidence of an individual having a deficit of at least –1.0 SD in 1 or more cognitive domains at study entry, dependent in part on the expected target specificity of the agent being tested. Although this cutoff might exclude a group of patients who may show some benefit to cognitive-enhancing agents, it will target those with the greatest need for treatment. Recent empirical approaches to subgrouping bipolar patients based upon neurocognitive profiles suggest that up to two-thirds of bipolar patients would meet the entry criterion of −1.0 SD on more than 1 domain, supporting this as a feasible cutoff.

Sample Size

Because of the clinical heterogeneity of bipolar disorder, cognitive outcome trials will require either large representative samples or smaller samples that are more narrowly defined with regard to particular clinical characteristics. Many clinical trials designed to assess the effects of pharmacotherapy on cognition in schizophrenia have included samples too small to detect a medium (d = 0.5) effect size. Based on MATRICS recommendations for pharmacologic studies, a randomized, double-blind, placebo-controlled, add-on design is optimal. Cross-over designs are not recommended for studies of cognition because of potential confounding of practice effects with treatment effects, as well as carryover effects of medications. It is worth noting, however, that the basic study design and many of the methodological strategies chosen will differ based upon the phase of study of the agent being investigated. This is particularly relevant when considering issues related to sample size and statistical power because early phase trials designed to detect a preliminary signal differ from pivotal trials in several ways. Proof-of-concept studies and phase 3 trials are crucial in drug development, particularly in psychiatrically vulnerable individuals, with the focus of initial studies being safety while still including efficacy measures in an effort to detect preliminary efficacy signals and to estimate effect sizes. This is especially true for agents that might be deemed high risk. For example, the likelihood that a medication might induce mania or exacerbate psychosis should be addressed before large-scale efficacy studies are launched.

Measurement of Primary Outcomes

Neurocognitive outcome measures. An important first step in the MATRICS initiative was the formation of a consensus battery to be uniformly applied to subsequent, large-scale clinical trials targeting cognitive dysfunction associated with schizophrenia. The MATRICS Consensus Cognitive Battery (MCCB) assesses the domains of attention/vigilance, speed of processing, working memory, verbal learning, visual learning, reasoning and problem solving, and social cognition. Although the MCCB was designed for use in cognitive trials of schizophrenia, its applicability to bipolar disorder has been tested. Burdick et al evaluated the MCCB in patients with bipolar I disorder and found significantly greater impairment relative to a matched control group on all domains except reasoning/problem solving and social cognition. Kessler et al reported similar results, showing significantly impaired performance on all MCCB domains in patients with both bipolar I and bipolar II disorders.

Unlike trials in schizophrenia, no consensus cognitive battery for bipolar disorder yet exists. Yatham et al convened a panel of neurocognitive and bipolar experts to begin to address the question of which tests might be optimal for use in patients with bipolar disorder; however, the recommended battery was not specific to clinical trials with cognition as an outcome but was considered by the panel for a broader application. Nonetheless, recommendations included the use of the MCCB as core, with additional measures added to increase sensitivity and specificity to bipolar disorder. The additional measures included the Stroop Test, the Trail...
Assessing Cognitive Function in Bipolar Disorder

Making Test-part B, the Wisconsin Card Sorting Test, and the California Verbal Learning Test (as an alternative to the Hopkins Verbal Learning Test-Revised used in the MCCB). Although this is only 1 potential battery, it addresses some of the critical issues specific to bipolar disorder. Because the deficits in bipolar disorder are less severe than those reported in schizophrenia, higher test sensitivity is important and could be achieved by including either more difficult measures or a greater number of tests tapping each domain (eg, several of the MCCB domain scores are derived from only a single measure). In addition, neurocognitive tests of affect, such as the affective Stroop Test or facial emotion recognition, should be considered when designing cognitive trials given their link to the pathophysiology of bipolar disorder. One existing schizophrenia battery (Brief Assessment of Cognition in Schizophrenia [BAC-S]) has been modified for use in affective disorders (Brief Assessment of Cognition in Affective Disorders [BAC-A]) and has been applied in a large-scale treatment trial of bipolar depression.

The choice regarding a primary outcome measure for cognitive trials in bipolar disorder will depend, at least in part, on the agent being studied and its purported mechanism of action; however, the use of a global cognitive measure, as is commonly used in schizophrenia trials, may be necessary if an FDA indication is being sought. Indeed, the best possible result is a drug that will improve each cognitive domain enough (or at least in the same direction) that it will improve global composite scores; however, it is crucial to highlight the lack of a global deficit in patients with bipolar disorder, which may require a different approach from that recommended in schizophrenia. Key secondary cognitive outcome measures will be important if a global composite is used as primary outcome and should be based upon a bipolar-specific profile, therefore focusing on tasks that measure attention, verbal memory, executive functions, and affective processing.

The recommended frequency of assessment will also depend upon the duration of a trial; however, testing at baseline and at least 2 time points thereafter is suggested to allow for potential attrition without total loss of data. In an 8-week study, for example, testing would be advisable at baseline, at 4 weeks, and again at study completion. Practice effects are nonnegligible in cognitive trials, and, although it is unclear how large practice effects will be among patients with bipolar disorder, they are likely to be comparable to practice effects noted in healthy subjects. The MCCB was designed to minimize these practice effects. Indeed, in a phase 3 trial of clinically stable outpatients with schizophrenia, the practice effect was small for the MCCB composite score. To account for practice effects, a placebo-controlled design is optimal; however, for early phase, open-label studies, estimated practice effects can be used if necessary based upon normative data.

Functional capacity outcome measures. The selection of appropriate measure(s) of functional outcome is crucial given the FDA stipulation that cognitive enhancement trials must show evidence of benefit on measures of functional capacity as well as objective cognitive assessment. The MATRICS guidelines note that coprimary measures should be included in cognitive trials and “should assess a clinically meaningful/relevant functional outcome, but not necessarily community functioning . . . [which] is highly dependent on psychosocial services, patient skills, and social support—factors that are usually beyond the control of clinical trials.” The University of California at San Diego Performance-Based Skills Assessment (UPSA) is the most widely used functional capacity measure in schizophrenia and has already been adopted for several trials in schizophrenia as a coprimary measure for FDA purposes. The UPSA has been assessed with patients with bipolar I disorder, and scores were significantly associated with indicators of impairment in everyday functioning; however, given the relatively lower level of functional disability in bipolar disorder than in schizophrenia, future studies must consider possible ceiling and practice effects on the brief version of the UPSA. Other widely used self-report measures that have been used in bipolar disorder affective efficacy trials, such as the Sheehan Disability Scale, might also prove useful and be efficient and cost effective. One measure, the Functioning Assessment Short Test (FAST), has been successful in tracking change in everyday functioning in patients with bipolar illness in nonpharmacologic remediation trials and may represent a promising tool.

Insomuch as relationships between neurocognitive performance and functional outcome are similar in schizophrenia and bipolar disorder, overlapping approaches could be considered appropriate for both patient groups. Emerging evidence does, indeed, suggest that the associations between cognitive impairment and functional capacity in schizophrenia pertain in a similar manner to cohorts of patients with bipolar disorder. Consistent with trial design in schizophrenia, cognitive trials in patients with bipolar disorder should incorporate measures of both functional capacity (having the skills necessary to perform an activity, eg, what a person can do) and everyday functioning in the community (eg, what a person does do, which may be influenced by even subthreshold depressive symptoms). In shorter-duration trials, functional capacity measures (eg, the UPSA) might be more tractable and amenable to change than are assays of how a person is functioning in the community (eg, employment status). Although the MATRICS committee recommends a study duration of at least 6 months for phase 3 trials to demonstrate an enduring effect on cognition in schizophrenia, this may be less feasible in bipolar patients because of the need to control for normal cycling associated with the illness. Finally, the use of objective measures of everyday functioning that utilize ratings made by an informant is generally recommended because patients’ self-reports of functioning are often inaccurate.

SUMMARY

Neurocognitive performance is significantly impaired in a substantial proportion of patients with bipolar disorder, despite adequate control of affective symptoms. These deficits are strongly associated with disabilities in community
function\textsuperscript{24,25} and warrant treatment.\textsuperscript{49} In contrast to a concerted effort to target this domain in patients with schizophrenia (eg, MATRICS), very few trials have been conducted in bipolar disorder. As the treatment of bipolar disorder evolves from the initial focus on controlling primary affective symptoms to the current consideration of secondary targets for a more complete recovery, it will be important to address the challenges inherent in this complex disorder. In this article, we have highlighted several illness-specific confounders that arise when designing cognitive trials in bipolar disorder, and we have made several suggestions for handling these a priori (Table 1). While some of the considerations discussed regarding cognitive trials are unique to these types of outcome measures, several issues are more broadly relevant to the design of experimental treatment trials in bipolar disorder, even when targeting the primary affective symptoms.\textsuperscript{65} Here, we propose a design that highlights the importance of both internal and external validity but ensures the potential for improvement by requiring sufficient cognitive impairment at the time of study entry by limiting enrollment to those patients having at least \textasciitilde 1.0 SD in 1 or more cognitive domains at study entry. Although the field remains in its infancy, a consensus on approaches to be implemented will be critical for facilitating advances in this area. It will be important to systematically address challenges as they arise during the conduct of future studies and to work as a field to optimize the chances of positive study outcomes, with the ultimate goal of optimizing the chances of positive patient outcomes. We present here some preliminary suggestions based upon the available evidence; however, the time is upon us to consider a more comprehensive, formalized consensus effort parallel to that conducted in the field for cognitive trials in schizophrenia (eg, MATRICS).

**Drug names:** aripiprazole (Abilify), bupropion (Wellbutrin, Aplenzin, and others), carbamazepine (Carbatrol, Equetro, and others), clozapine (Clozaril, FazaClO, and others), fluoxetine (Prozac and others), lamotrigine (Lamictal and others), lithium (Lithobid and others), luraspidine (Latuda), olanzapine (Zyprea and others), paroxetine (Paxil, Paxeva, and others), pramipexole (Mirapex and others), quetiapine (Seroquel and others), sertraline (Zoloft and others), topiramate (Topamax and others), ziprasidone (Geodon and others).

**Author affiliations:** Departments of Psychiatry (Drs Burdick and Goldberg) and Neurosciences (Dr Burdick), Icahn School of Medicine at Mount Sinai, New York, New York; Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California (Dr Ketter); and Department of Psychiatry, Case Western Reserve University School of Medicine, Cleveland, Ohio (Dr Calabrese).

**Potential conflicts of interest:** Dr Burdick has served as an advisory board member for Dainippon Sumitomo and Takeda Lundbeck. Dr Ketter has served as a consultant to Genentech, Janssen, Merck, Prophase, Sunovion and Teva; has received grants from Agency for Healthcare Research and Quality and from Sunovion; has received honoraria from Pfizer; and has received royalties from American Psychiatric Publishing. His spouse is employed by and owns stock in Janssen. Dr Goldberg has served as a consultant, advisory board member, or speakers’ bureau member for AstraZeneca, Avanir, Medscape, Sunovion, Takeda Lundbeck, and WebMD; receives royalties from American Psychiatric Publishing; and has received material support from Frontline Medical Communications. Dr Calabrese has received research support from Pfizer, AstraZeneca, & Cephalon; has been a lecturer supported by American Foundation Suicide Prevention, AstraZeneca, Benecke, CME Outfitters, Dainippon Sumitomo Pharma, Elan, Forest, Health & Wellness Partners, Lundbeck, Medviz, Otsuka, ProMedica, Spirant Communication Private Limited, Sunovion, Takeda, Teva, and Wencsekach Institute; and has consulted for Biomedical Development Corporation, Convergent Health Solutions, Dainippon Sumitomo Pharma, Elan, Forest, Health & Wellness Partners, Hoffman LaRoche, Lilly, Lundbeck, Otsuka, Scientia, Sunovion, Takeda, and Teva.

**Funding/support:** None reported.

**Acknowledgments:** Under the direction of the authors, a literature search was conducted and an outline drafted by Nancyolland, PhD, and Patricia Johansen, PhD, Synchrony Medical Communications, LLC, West Chester, Pennsylvania. Drs Holland and Johansen received funding for this work from Teva Pharmaceuticals, North Wales, Pennsylvania.

**REFERENCES**


