The FDA-NIMH-MATRICS Guidelines for Clinical Trial Design of Cognitive-Enhancing Drugs: What Do We Know 5 Years Later?

Robert W. Buchanan*,1, Richard S. E. Keefe2, Daniel Umbricht3, Michael F. Green4,5, Thomas Laughren6, and Stephen R. Marder4,5

1Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, PO Box 21247, Baltimore, MD 21228; 2Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC 27710; 3F. Hoffmann-La Roche Ltd, PDEN—CRED Neuroscience, 4070 Basel, Switzerland; 4University of California at Los Angeles Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA 90095; 5Veterans Administration Veterans Integrated Service Network, 22 Mental Illness Research, Education, and Clinical Center, Los Angeles, CA 90073; 6Psychiatry Products Division, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD 20993

*To whom correspondence should be addressed; tel: 410-402-7876, fax: 410-402-7198, e-mail: rwbuchanan@mprc.umaryland.edu.

The Food and Drug Administration (FDA)-National Institute of Mental Health (NIMH)-Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) clinical trial guidelines for cognitive-enhancing drugs in schizophrenia and the MATRICS Consensus Cognitive Battery (MCCB) were designed to facilitate novel compound development in the treatment of cognitive impairments. Several studies have recently utilized the FDA-NIMH-MATRICS guidelines and MCCB and allow an evaluation of the feasibility of guideline implementation and MCCB performance. In light of the study results, we would recommend the following inclusion criteria revisions—(1) clinical status and symptom inclusion criteria: maximum allowed score for hallucinations and delusions should be increased from moderate to moderately severe and the negative symptom criterion should be dropped in phase 2 studies; (2) antipsychotic medication inclusion criteria: first-generation antipsychotics should be allowed, but only in the context of no concomitant anticholinergic agents and minimal extrapyramidal symptoms, and antipsychotic polypharmacy should be allowed in the absence of pertinent pharmacokinetic or pharmacodynamic considerations; and (3) people who use illicit substances should not be allowed in phase 1B or 2A proof-of-concept studies but may be included in phase 2B and 3 studies in which proof of effectiveness and generalizability of results become more important goals. These revisions are recommended to enhance recruitment while maintaining sufficient methodological rigor to ensure the validity of study results. The MCCB has been shown to have excellent psychometric characteristics, including reliability for multisite clinical trials, clinical relevance for real-world functioning, and possible sensitivity to behavioral treatment, and should continue to serve as the standard outcome measure for cognitive enhancement studies in schizophrenia.

Key words: study design/drug development/cognitive impairments/substance abuse/neuropsychological test battery

Introduction

In April 2004, Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) investigators organized an US Food and Drug Administration (FDA)/National Institute of Mental Health (NIMH) consensus meeting to develop guidelines for the design of clinical trials of cognitive-enhancing drugs for people with schizophrenia.1 Participants included representatives from government (FDA and NIMH), academia, and industry. The workshop developed recommendations for subject selection, co-primary outcome measures, and statistical approaches for study design. The guidelines were conceptualized as a reasonable starting point for use in trial design of cognitive-enhancing drugs. In the last 5 years, several studies have been conducted using the FDA-NIMH-MATRICS guidelines and the MATRICS Consensus Cognitive Battery (MCCB).2,3 These trials provide new information on the feasibility and relevance of the proposed guidelines and MCCB.

In the present article, we (1) re-evaluate those guidelines for which there is new information; (2) propose new criteria for handling the problem of illicit substance use among trial participants; (3) evaluate the performance of the MCCB in 2 industry-sponsored, multisite clinical trials; and (4) update the status of the evaluation of co-primary measures.

© The Author 2010. Published by Oxford University Press on behalf of the Maryland Psychiatric Research Center. All rights reserved. For permissions, please email: journals.permissions@oxfordjournals.org.
Inclusion Criteria

MATRICS investigators evaluated the feasibility of implementing studies based on the FDA-NIMH-MATRICS guidelines through the design and conduct of 2 Treatment Units for Research on Neuropsychology and Schizophrenia (TURNS) clinical trials. The first study (AL-108 study) was a 12-week, double-blind, placebo-controlled randomized clinical trial of AL-108, an 8 amino acid peptide fragment of activity-dependent neuroprotective protein (http://clinicaltrials.gov/, trial number: NCT00505765). Two doses of AL-108, 5 mg/day and 15 mg twice daily (BID), were compared with placebo. The second study (MK-0777 study) was a 4-week, double-blind, placebo-controlled randomized clinical trial of MK-0777, a gamma-amino butyric acidA alpha2/3 partial agonist (http://clinicaltrials.gov/, trial number: NCT00505076). Two doses of MK-0777, 3 mg BID and 8 mg BID, were compared with placebo. The FDA-NIMH-MATRICS guidelines propose inclusion criteria for the diagnosis; clinical state, including symptom severity; antipsychotic medications; concomitant medications; and level of cognitive impairment. We followed the proposed inclusion criteria guidelines for diagnosis, concomitant medications, and level of cognitive impairment. The 2 areas in which we deviated from the original guidelines were the criteria used to define the clinical state of potential participants and the criteria used to define allowed antipsychotic treatment.

Clinical State and Symptom Severity

The intent of the clinical state and symptom severity inclusion criteria is to facilitate the isolation of a specific effect on cognition from other concurrent changes in clinical status that may secondarily affect cognitive function, eg, an improvement in positive symptoms may improve cognitive measures because of an enhanced ability to cooperate and focus on the cognitive tasks rather than through a primary effect on cognition. In order to achieve this goal, the FDA-NIMH-MATRICS guidelines propose that potential participants should (1) be clinically stable and in the residual (nonacute) phase of their illness for a specified period of time (at least 2 months); (2) be maintained on current antipsychotic and concomitant medications for a specified period of time (at least 6 weeks), with no change in antipsychotic dose for the previous 2–4 weeks; and (3) exhibit no more than moderate levels of hallucinations and delusions, positive formal thought disorder, and negative symptoms, with a minimal level of depressive and extrapyramidal symptoms (EPS). The rationale for these recommendations is that in combination, they would identify a group of potential participants whose clinical status, including symptoms, would be least likely to change during the course of a clinical trial. Previous studies suggest that change in clinical status is most likely to significantly affect neuropsychological test performance when there is a marked change in clinical status, eg, the comparison of an acutely psychotic, unmedicated baseline to a posttreatment endpoint. For review, see Buchanan et al and Keefe et al, Harvey et al, McEvoy et al, and Kahn.

In the 2 TURNS studies, the following changes were made in these criteria: (1) we increased the allowed maximum severity of hallucinations and delusions from moderate to moderately severe (ie, on the Brief Psychiatric Rating Scale [BPRS], item range 1–7, the maximum allowed item score was increased from 4 to 5) and (2) we dropped the restriction on negative symptom severity (see Table 1 for summary of proposed guideline changes). We did not change the requirement that positive formal thought disorder be of moderate severity or less.

We made these changes for the following reasons: (1) the original intent of the FDA-NIMH-MATRICS guidelines was to emphasize the requirement of clinical stability. In this context, the assurance that participants are optimally treated and clinically stable and in the residual (nonacute) phase of their illness is more important than the absolute level of positive and negative symptom severity and (2) the restriction on hallucination and delusion severity was potentially compromising our ability to recruit otherwise eligible subjects for the study. There are 3 potential concerns associated with these changes. The first is that allowing participants to enter the study with increased positive symptom severity may increase the risk of concurrent changes in positive symptoms and cognitive measures and runs the risk that any observed change in cognitive measures could be attributed to the changes in positive symptoms. However, the recent Clinical Antipsychotic Trial of Intervention Effectiveness study continues to suggest that, in the context of minimal change in clinical status, change in positive symptoms does not lead to change in neuropsychological test performance. The second concern is that increased hallucination, delusion, and negative symptom severity could potentially compromise the ability of the subject to validly complete the neuropsychological tests. There is substantial evidence from studies conducted in unmedicated participants that neuropsychological assessments can be validly completed in the presence of marked positive and negative symptoms. Furthermore, most studies have test validity criteria in place to ensure the quality of the neuropsychological test data. In contrast, there are at least 2 important practical reasons to retain the current positive formal thought disorder symptom severity guideline. In addition to previous studies suggesting that severity of positive formal thought disorder is associated with neuropsychological test performance, for review, see Buchanan et al, the presence of marked positive formal thought disorder may (a) compromise the ability to confirm that the participant has adequately understood test directions and (b) interfere with performance on language-based neuropsychological tests. Therefore, we did not change this criterion.
The third concern is related to the lack of a negative symptom criterion. The relationship between cognitive impairments and negative symptoms in schizophrenia is unclear. Although a common or overlapping pathophysiology has been hypothesized, the extent to which there is a shared pathophysiology has not been clarified. In light of the unknown relationship between these 2 domains, the restriction of potential participants to those with cognitive impairments in the absence of negative symptoms may actually compromise the development of cognitive-enhancing agents because such a study population may represent a unique subset of people with schizophrenia and cognitive impairments. In addition, there is the possibility that the absence of a negative symptom criterion could lead to concurrent changes in both cognition and negative symptoms. In view of the fact that these 2 domains represent the 2 major unmet therapeutic areas in schizophrenia and they are the 2 most important predictors of functional impairments, the development of a drug that improved both cognitive function and negative symptoms would be of considerable clinical benefit. However, if future trials reveal that improvements in cognition are always accompanied by improvements in negative symptoms, then there will be a need to rethink how to characterize this broader construct of residual phase impairment, especially with regard to labeling claims. We did not change either the depressive or EPS criteria because neither had a practical impact on recruitment (S. R. Marder et al, unpublished data from the AL-108 and MK-0777 studies).

**Antipsychotic Treatment**

The other major change in the previous inclusion criteria was related to antipsychotic treatment (see Table 1 for summary of proposed guideline change). The intent of the originally proposed criteria was to minimize pharmacokinetic or pharmacodynamic interactions that could compromise the evaluation of experimental drug efficacy. In order to achieve this goal, the FDA-NIMH-MATRICS guidelines proposed that the selection of allowed antipsychotics should consider potential interactions of the experimental agent with the known pharmacological properties of the various antipsychotics (eg, the use of high-dose olanzapine might not be allowed in a study of a nicotinic agent) and should prohibit antipsychotic polypharmacy. There was no specific prohibition of first-generation antipsychotics (FGAs), but the guidelines noted that FGAs were more likely than second-generation antipsychotics (SGAs) to cause EPS and were more likely to be used in conjunction with anticholinergic agents, which can compromise memory function. The combination of these 2 factors certainly raises concerns about allowing the use of FGAs.

In the context of the 2 TURNS studies, we made the following changes in the antipsychotic treatment criteria.

**Table 1. Recommended Modifications of Original FDA-NIMH-MATRICS Guidelines**

<table>
<thead>
<tr>
<th>Original Guideline</th>
<th>New Guideline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To isolate change in cognitive function from change in symptoms and other clinical features, include subjects who (a) have no more than a “moderate” severity rating on hallucinations and delusions (eg, BPRS hallucinatory behavior or unusual thought content item score ≤4), (b) have no more than a moderate severity rating on positive formal thought disorder (eg, BPRS conceptual disorganization item score ≤4), and (c) have no more than moderate severity rating on negative symptoms (eg, all Scale for the Assessment of Negative Symptoms global items &lt;3 or PANSS negative syndrome total score ≤15)</td>
<td>1. To isolate change in cognitive function from change in symptoms and other clinical features, include subjects who (a) have no more than a moderate severity rating on hallucinations and delusions (eg, BPRS hallucinatory behavior or unusual thought content item score ≤5) and (b) have no more than a moderate severity rating on positive formal thought disorder (eg, BPRS conceptual disorganization item score ≤4).</td>
</tr>
<tr>
<td>2. Exclude participants who are taking more than 1 antipsychotic. Antipsychotic combinations are an unnecessary complication in studying an adjunctive/cotreatment agent</td>
<td>2. In the absence of significant pharmacokinetic or pharmacodynamic interactions between the antipsychotic combination and the experimental drug, up to 2 antipsychotics are permitted.</td>
</tr>
<tr>
<td>3. Not applicable</td>
<td>3. In the absence of significant pharmacokinetic or pharmacodynamic interactions and concomitant anticholinergics and in the presence of low extrapyramidal symptoms, first-generation antipsychotics are permitted.</td>
</tr>
<tr>
<td>4. Not applicable</td>
<td>4. Exclude participants who use illicit substances but do not meet DSM-IV-TR criteria for substance abuse or dependence in phase 1B and 2A studies but include in phase 2B and 3 studies.</td>
</tr>
</tbody>
</table>

*Note: BPRS, Brief Psychiatric Rating Scale; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; PANSS, Positive and Negative Syndrome Scale; TR, Text Revision.*
In the AL-108 study, we allowed the use of FGA long-acting injectable preparations when used in the absence of concomitant anticholinergics. The criterion for no dose change was extended to the last 3 months for these agents. The rationale for this change was 2-fold. First, several recent studies suggest that there are minimal FGA/SGA differences in their effect on cognitive function.\(^4,8,20,21\) Second, the continued use of the EPS criteria and the continued restriction on anticholinergics further minimizes the differential neurotoxic effects of FGAs and SGAs that are observed when high doses of high-potency FGAs are compared with SGAs.

In the MK-0777 study, we allowed antipsychotic polypharmacy because of concerns that the prohibition of antipsychotic polypharmacy could substantially impact recruitment. The prevalence of antipsychotic polypharmacy varies across study population but ranges from approximately 10% to 37%.\(^22–25\) In the absence of significant pharmacokinetic or pharmacodynamic considerations related to a particular antipsychotic combination with relevance to the experimental drug under consideration, there is no compelling reason to preclude antipsychotic polypharmacy.

### Illicit Substance Abuse/Dependence

The FDA-NIMH-MATRICS guidelines propose inclusion criteria for a number of diagnostic, clinical, and treatment variables but are silent on the issue of illicit substance use. A recent proof-of-concept trial of R3487 cosponsored by Memory Pharmaceuticals Inc and F. Hoffmann-La Roche, Ltd, provides critical new information on this issue (http://clinicaltrials.gov/, trial number: NCT00604760). The study followed closely the original FDA-NIMH-MATRICS guidelines\(^1\): participants were judged to be clinically stable and in the nonacute phase of their illness for at least 12 weeks; have a baseline Positive and Negative Syndrome Scale (PANSS) total score of 70 or less; no more than a “moderate” severity rating on the PANSS hallucinatory behavior item; and a PANSS positive syndrome subscale score of 20 or less. There was no restriction on negative symptom severity. In addition, participants who met Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) diagnosis of alcohol or substance abuse or dependence (other than nicotine) within the last 6 months were excluded. Participants were also required to have a negative baseline urine drug screen. In the original version of the protocol, 2 consecutive positive urine drug screens would lead to participant withdrawal. The illicit drug use exclusion criterion was included because of the concern that such use could affect the validity of the assessment of cognition.

A total of 215 participants were randomized. In contrast to expectations, many more participants than anticipated had positive urine drug screens for illicit drugs. The number of positive drug screens was so large that the completion of the study was actually jeopardized. Among all randomized participants, 20% had a positive urine drug screen during the trial. More importantly, among those completing the trial, 31% of the participants had a positive drug screen. Because the adherence to a strict standard of excluding participants with evidence of recent illicit substance use would have led to a substantial reduction in sample size and power to detect potential treatment effects, the withdrawal criterion was eliminated and the decision to include or exclude the participant with a positive drug screen was left to the clinical judgment of the site investigator.

In light of our experience with this proof-of-concept trial, we would propose that the issue of illicit substance use should be given due consideration, with the study design procedures used to deal with this issue dependent upon the particular phase of drug development (see Table 1 for proposed guideline). The strict exclusion of participants who occasionally use illicit drugs is designed to reduce potential pharmacological confounding effects but may substantially affect the recruitment and retention of participants and the generalizability of study results. In the context of proof-of-concept studies, the primary focus is the initial confirmation of a positive effect on cognition and the issue of generalizability is of less concern. Hence, reduction of potential confounding factors may be more important than generalizability, and participants with occasional, illicit substance use should be excluded. However, once proof of concept has been established, the evaluation of a novel compound in real-world populations and generalizability becomes more important and the exclusion of participants who occasionally use illicit substances needs to be weighed against concerns about generalizability, recruitment, and retention. We would recommend the inclusion of such participants, with stratification based on baseline illicit substance use and post hoc sensitivity analyses that control for the potential impact of participants whose illicit substance use is detected after randomization. In addition, we would strongly recommend that any participant who used illicit substances immediately prior to any testing occasion be rescheduled.

### MATRICS Consensus Cognitive Battery

The MATRICS initiative produced a battery of tests, the MCCB, to assess cognitive treatment effects in people with schizophrenia. A set of norms has been published,\(^3\) a manual is available,\(^26\) and all component measures are available in a bundled kit. In validation studies, the MCCB demonstrated excellent reliability, minimal practice effects, and significant correlations with measures of functional capacity.\(^2\) A major empirical question is whether the MCCB would demonstrate these same favorable characteristics when administered in the context of large multisite industry trials for which it was designed.
Two recently completed industry-sponsored, multisite cognition trials have followed the FDA-NIMH-MATRICS guidelines and used the MCCB as a primary outcome measure. These studies provide excellent data for examining the MCCB as a repeated measure and will be used to examine the following psychometric characteristics of the MCCB: test-retest reliability, practice effects, placebo effects, and correlations with functional capacity.

Both studies used the MCCB composite score as the primary outcome measure with 5 assessment points, including 2 assessments prior to the initiation of treatment, which allows a real-world examination of test-retest reliability characteristics. The first study was the above-referenced proof-of-concept trial of R3487 cosponsored by Memory Pharmaceuticals Inc and F. Hoffmann-La Roche, Ltd (http://clinicaltrials.gov/, trial number: NCT00604760). Two hundred fifteen clinically stable outpatients were enrolled from 22 US sites.27 The MCCB was used to assess cognitive functioning at screening, baseline, and weeks 4, 8, and 10. The MCCB generates 7 cognitive domain scores and a composite score that sums and standardizes data from all 7 domains based upon published normative data.3 Standardized T scores (mean of $T = 50$ and SD = 10) are generated for each domain and the overall composite score. In order for an assessment to be considered sufficiently complete to generate a composite score, data from 5 of 7 domains must be available.26 In this study, functional capacity was assessed with the University of California at San Diego Performance-based Skills Assessment, 2nd edition (UPSA-2; unpublished revision of the UPSA), at baseline.

The MCCB had a very low rate of missing data with only 14 of 4300 measures missing and none of the 215 study participants had a missing composite score. The screening and baseline test data were used to evaluate the MCCB composite and domain score test-retest reliability and short-term practice effects. The interval between screening and baseline was 1–2 weeks. The MCCB composite score had excellent test-retest reliability (intrarater correlation coefficient [ICC] = 0.88) and sensitivity to impairment at screening ($T$ score = 25.1 ± SD = 11.6) and baseline ($T$ score = 26.7 ± 12.1). The test-retest reliabilities of the individual domain scores were not as high, though the correlations ranged between good ($ICC = 0.70$) and excellent ($ICC = 0.80$) for most domains. The only exception was verbal learning ($ICC = 0.65$), in which alternate forms of the Hopkins Verbal Learning Test were used at each assessment. $Z$ scores were used to assess practice effects for the MCCB composite score and for each domain score. The $z$ scores were calculated by subtracting the mean $T$ score of the measure at screening from the mean $T$ score at baseline and then dividing the difference score by the SD of the measure at the screening visit. The practice effect for the MCCB composite score ($z = 0.20$) and each of the 7 domains was small ($z = 0.01–0.27$). The Pearson correlation between the baseline MCCB composite and UPSA-2 total scores was 0.56 ($P < .001$), confirming the construct validity of the MCCB and the UPSA. The practice effects over an extended period were evaluated using the placebo group baseline, week 4, and week 8 MCCB data. The effect size of the MCCB composite score improvement over this time period was small ($z = 0.2$). The magnitude of this effect is consistent with the practice effect observed in the total sample comparison of screening and baseline performance data and suggests that repeated administration of the MCCB is associated with minimal practice effects. The reliability between assessments in the placebo group was high (all ICCs >0.90).

The second trial confirmed these findings (http://clinicaltrials.gov/, trial number: NCT00641745).29 This trial used the MCCB to assess cognition at screening, baseline, week 6, month 3, and month 6. There were 323 participants from 29 US sites. The interval between screening and baseline assessments was more varied, with a median interval of 15 days. Functional capacity was assessed with the UPSA—Brief Version (UPSA-B).30 The severity of cognitive impairment at screening ($T$ score = 24.7 ± SD = 12.1) and baseline ($T$ score = 26.9 ± SD = 12.4) was similar to the first trial. The screening to baseline test-retest reliability of the MCCB composite score was identical to the first trial (ICC = 0.88), and the test-retest reliabilities of the domain scores were also very similar. Only 14 test scores were missing out of a total of 6460 test assessments (10 MCCB assessments performed in 323 subjects at 2 occasions; 99.8% complete). At baseline, all 323 participants met MCCB criteria for sufficient data to compute a composite score. Construct validity was again robust because the baseline MCCB composite score was significantly correlated with the baseline UPSA-B composite score ($r = 0.61$, $df = 304$, $P < .001$). The MCCB composite score practice effect was small ($z = 0.18$).

These results suggest that the MCCB has proven reliability for multisite clinical trials and clinical relevance for real-world functioning and should continue to serve as the standard for cognitive enhancement studies in schizophrenia. However, these data do not address the issue of whether the MCCB is sensitive to drug treatment-related change. The lack of compounds with proven efficacy for cognitive impairment in schizophrenia precludes the direct assessment of the sensitivity of the MCCB to pharmacological treatment. To date, there are 2 published studies that have assessed the efficacy of experimental pharmacological agents on the MCCB composite score. Neither of these studies demonstrated a significant beneficial effect, but both were nondefinitive owing to either small sample sizes31 and/or crossover designs that confounded practice effects with treatment effects.14 Several large-scale studies using the MCCB are currently underway (http://clinicaltrials.gov/, accessed February 2010), and a wealth of data on the sensitivity of the battery to treatment changes will be available in the next year.
or 2. However, recent data suggest that the MCCB composite score is sensitive to a behavioral intervention, with an effect size of $d = 0.88$ for the PositScience cognitive remediation program compared with a control program of standard video games. 32 While the effects of behavioral intervention may differ from the effects of pharmacological intervention, these behavioral data support the notion that the MCCB has the potential to be sensitive to cognitive improvement with pharmacological treatment.

Co-primary Measures

The FDA has emphasized at numerous public meetings that significant improvement in a cognitive performance endpoint would be a necessary, but not sufficient, condition for drug approval. In addition to the required changes in cognitive performance, it will be necessary to demonstrate improvement on a co-primary functional measure that would have greater face validity to consumers and clinicians than cognitive performance measures. 1

The requirement for a co-primary functional measure presents a challenge because there is no consensus on a particular measure for this purpose or even a consensus on what type of measure is the most appropriate to use in clinical trials. Ideally, one would want an assessment of community functioning as a co-primary measure. However, changes in community status (eg, work, social relationships, or degree of independent living) are likely to be too stringent a requirement because it would involve multiple intervening variables (eg, psychosocial rehabilitation, family support, local employment rates, etc.) that act between underlying cognitive processes and functional outcomes. Such intervening factors are typically outside the control of a clinical trial and would obscure the expected functional benefits of cognition-enhancing effects.

The MATRICS Psychometric and Standardization Study (PASS) evaluated potential co-primary measures derived from 2 different approaches: performance-based tests of functional capacity and interview-based assessments of cognition. 33 The assessment of functional capacity involves simulation of real-world activities, such as holding a social conversation, preparing a meal, or taking public transportation. 34, 35 These measures assess whether a person can perform the task when given an opportunity, but intact performance on these measures does not guarantee that a person will perform these tasks in the community. Because performance on functional capacity measures does not depend on social and community opportunities, they are more likely to be temporally linked to treatment-related changes in cognition. The second approach, interview-based assessments of cognitive abilities, involves asking people to estimate their own cognitive abilities and to estimate the extent to which their daily lives are affected by cognitive impairment.

Several cognitive assessment interviews have been developed for people with psychotic disorders. 36, 37 In a sample of people with schizophrenia, the MATRICS PASS evaluated 2 measures from each type of approach. In brief, the main study conclusions were as follows: (1) all 4 of the measures had acceptable test-retest reliability, (2) the relationships to cognitive performance were higher for functional capacity measures than for the interview-based measures of cognition, (3) all the measures had comparably modest relationships to community functional status, and (4) the interview-based measures had more missing data due to difficulty in finding suitable informants. 33

The MATRICS PASS was an initial attempt to evaluate co-primary measures, and it serves as a guide for subsequent efforts. Because the need for co-primary measures was not known at the start of the MATRICS initiative, the selection of potential co-primary measures did not go through a rigorous selection process similar to the selection of the MCCB measures (eg, nominations, review of the literature, a formal RAND Panel, etc.). 38 Pharmaceutical companies wanted more empirical guidance and a clearer consensus on which co-primary measures to use in cognition enhancement studies in schizophrenia. To address this need, a new government-industry-academic consortium was formed: MATRICS Co-primary and Translation (MATRICS-CT). 39 At present, the MATRICS-CT consortium includes 14 pharmaceutical companies and is coordinated by the Foundation for the National Institutes of Health. As the name implies, the consortium was formed to address 2 unresolved issues from the MATRICS initiative: (1) the lack of consensus on co-primary measures to be used in clinical trials and (2) the need to translate the MCCB into foreign languages for international clinical trials.

MATRICS-CT has launched a series of steps to evaluate co-primary measures that were modeled on those used to evaluate cognitive performance measures. 38 First, selection criteria for the measures were suggested and presented at a public meeting in August 2007. Second, nominations were solicited for potential co-primary measures. The measures that were nominated tended to fall into the same 2 approaches that were used in the MATRICS PASS: performance-based functional capacity measures and interview-based assessments of cognition. Third, the nominations were evaluated, and a subset of the nominated measures was selected for further consideration by a RAND Panel. Fourth, to prepare for the RAND Panel, a database was compiled of all available published and unpublished data that were relevant for the selection criteria. Fifth, a RAND Panel was held in February 2008 to evaluate the proposed measures. Sixth, co-primary measures were selected based on the RAND Panel ratings for inclusion in an empirical study. This study, called the Validation of Intermediate Measures (VIM) Study, evaluated the measures in 166 people.
with schizophrenia at 4 performance sites. The goal was to gather data on the psychometric characteristics of each measure, as well as validity criteria, such as relationships to cognitive performance and daily functioning. Each participant was tested twice, at a 4-week interval. The VIM Study included 3 performance-based measures of functional capacity: the Test of Adaptive Behavior in Schizophrenia (TABS), UPSA, and the Independent Living Scales, as well as 2 interview-based measures of cognition: the Cognition Assessment Interview (CAI) and the Clinical Global Impression Scale for Cognition. The data from the VIM Study were presented at a NIMH public meeting in October 2009 and information is available at: http://www.matrics.ucla.edu/matrics-ct/home.html. Finally, a survey of mental health researchers in representative countries is being conducted to obtain opinions about degree to which the co-primary measures can be adapted for use in other countries.

The information from the MATRICS-CT consortium will be helpful to companies as they select co-primary measures for clinical trials. Candidate co-primary measures for functional capacity and interview-based measures of cognition are relatively early in development, and all the measures evaluated by MATRICS-CT have recently been revised. Hence, unlike the situation with the MCCB, in which potential measures were further along in their development, the FDA has not made a specific request to identify a single consensus co-primary measure.

Discussion

The primary goal of the FDA-NIMH-MATRICS guidelines for the design of clinical trials of cognitive-enhancing drugs in schizophrenia was to facilitate the development of new agents for this indication. The original FDA-NIMH-MATRICS guidelines were conceptualized as a reasonable starting point, with the stated understanding that new data would lead to guideline modifications. There have now been a series of studies that have examined the effect of various pharmacological interventions on cognitive performance, including several that have utilized the FDA-NIMH-MATRICS guidelines. In response to the new information from these studies, we have proposed the following changes—(1) clinical status and symptom inclusion criteria: the maximum allowed score for hallucinations and delusions should be decreased from moderate to moderately severe (e.g., BPRS or PANSS score of 5) and the negative symptom criterion should be dropped, especially for phase 2 studies, but also could be potentially dropped for phase 3 studies, and (2) antipsychotic medication inclusion criteria: FGAs should be allowed, but only in the context of no concomitant anticholinergic agents and the continued use of the EPS inclusion criterion, and antipsychotic polypharmacy should be allowed in the absence of pertinent pharmacokinetic or pharmacodynamic considerations. The experience to date with the FDA-NIMH-MATRICS guidelines does not suggest that there should be any changes made in the maximum allowed score for positive formal thought disorder or in the depressive and EPS and level of cognitive impairment criteria.

The issue of illicit substance use was not adequately addressed in the original FDA-NIMH-MATRICS guidelines. In light of the widespread use of illicit substances, the blanket exclusion of participants who use these substances could markedly undermine the ability to conduct registration studies and limit the generalizability of findings from such studies. Therefore, we would recommend that people who use illicit substances, but do not meet DSM-IV criteria for abuse or dependence, be allowed to participate in phase 2B and 3 studies, but not in phase 1B or 2A proof-of-concept studies, in which the primary goal is the detection of an efficacy signal.

The results of 2 large, industry-sponsored, multisite clinical trials have confirmed the positive psychometric characteristics of the MCCB observed in the initial validation studies. In particular, these studies found that the MCCB has very good test-retest reliability, minimal practice effects, and significant correlations with measures of functional capacity. The FDA continues to recommend the MCCB as the “gold standard” for the assessment of cognition in schizophrenia clinical trials. However, the FDA is open to the consideration of the use of alternative neuropsychological test batteries if the batteries are comprised of measures of the key schizophrenia cognitive domains identified through the MATRICS process and are demonstrated to be valid for assessing these domains. Ideally, these alternative instruments would also have psychometric characteristics equivalent to those of the MCCB and have normative data as well in the settings where they would be used.

Finally, there is a broad range of potential co-primary functional measures. There are several reasons why functional outcome measures may not be feasible to use, but performance-based functional capacity measures or interview-based assessments of cognition could be used for this purpose. The FDA would accept either type of these measures. The MATRICS-CT VIM Study results will provide information to guide the selection of specific co-primary assessments.

Funding

National Institute of Mental Health (contract MH22006): Measurement and Treatment Development Activities on Cognition in Schizophrenia (Marder, Green); National Institute of Mental Health (HHSN278200441003C): Treatment Units for Research on Neurocognition and Schizophrenia (Marder); Janssen Pharmaceutica to Buchanan; Allon, Astra-Zeneca, GlaxoSmithKline, Novartis, the
Singapore National Medical Research Council to Dr Keefe; Novartis, GlaxoSmithKline to Marder.

Acknowledgments

Dr Laughren is with the Food and Drug Administration (FDA), Silver Spring, MD. The contributions of Dr Laughren were made in his private capacity; no official support or endorsement by the FDA is intended or should be inferred; Dr Laughren has no competing interests or financial support to disclose. Buchanan—Data and Safety Monitoring Board member: Cephalon, Otsuka, and Pfizer; Consultant: Abbott, GlaxoSmithKline, Sanofi-Aventis, and Scherring-Plough; Advisory Boards: Abbott, Astra-Zeneca, Merck, Pfizer, Roche, Solvay Pharmaceuticals, Inc, and Wyeth; Dr Keefe—Advisory Board/Consultant: Abbott, Astra-Zeneca, BiolineRx, Bristol Myers Squibb, Cephalon, CHDI Foundation, Inc., Dainippon Sumitomo Pharma, Eli Lilly, EnVivo, Johnson & Johnson, Lundbeck, Memory Pharmaceuticals, Merck, Neurosearch, Orion, Orexigen, Otsuka, Pfizer, Roche, Sanofi-Aventis, Shire, Solvay, Takeda, Wyeth, and Xenoport. Dr Keefe receives royalties from the Brief Assessment of Cognition in Schizophrenia (BACS) testing battery and the MATRICS Battery (BACS Symbol Coding). Dr Keefe is the Founder of NeuroCog Trials, Inc, a company that provides rater training and data quality assurance for clinical trials that assess cognition with various measures, including the BACS and the MCCB; Dr Umbricht is an employee of F. Hoffmann-La Roche Ltd; Marder—Advisory Boards: Wyeth, Pfizer, Schering Plough, Sanofi-Aventis, Lundbeck, and Roche; Green—Consultant: Abbott Laboratories, Dainippon Sumitomo Pharma, Otsuka, Sanofi-Aventis, and Takeda; Speaker: Janssen Cilag. He is an officer in MATRICS Assessment Inc but receives no financial compensation for that role and no financial compensation from the sale of the MATRICS Consensus Cognitive Battery.

References


