

## A Summary of the FDA-NIMH-MATRICES Workshop on Clinical Trial Design for Neurocognitive Drugs for Schizophrenia.

**Robert W. Buchanan, Miriam Davis, Donald Goff, Michael F. Green, Richard S. E. Keefe, Andrew C. Leon, Keith H. Nuechterlein, Thomas Laughren, Robert Levin, Ellen Stover, Wayne Fenton, and Steve R. Marder**

Send reprint requests to Robert W. Buchanan, M.D., Maryland Psychiatric Research Center, P.O. Box 21247, Baltimore, MD 21228; e-mail: RWBuchanan@mprc.umaryland.edu.

**Objective:** On April 23, 2004, a joint meeting of the FDA, NIMH, MATRICES investigators, and experts from academia and the pharmaceutical industry was convened to develop guidelines for the design of clinical trials of cognitive-enhancing drugs for neurocognitive impairments in patients with schizophrenia. **Method:** Experts were asked to address specific questions relating to clinical trial design of adjunctive/co-treatment and broad spectrum agents. At the workshop, experts reviewed relevant evidence before offering the discussion panel proposed guidelines for a given subset of questions. The discussion panel, which consisted of presenters and representatives from FDA, NIMH, academia, and industry, deliberated to reach consensus on suggested guidelines. When evidence was insufficient, suggested guidelines represent the opinion of a cross-section of the presenters and discussion panel. **Results:** Guidelines were developed for inclusion criteria, the use of co-primary outcome measures, and statistical approaches for study design. Consensus was achieved regarding diagnostic and concomitant medication inclusion criteria and on the use of cognitive screening measures. A key guideline was to limit the trial to patients in the residual phase of their illness, who have a predefined level of positive, negative, and affective symptoms. The most difficult issues were the feasibility of including a co-primary measure of functional improvement and the choice of comparator agent for a trial of a broad spectrum agent (with antipsychotic and cognitive-enhancing effects). **Conclusions:** The suggested guidelines represent reasonable starting points for trial design of cognitive-enhancing drugs, with the understanding that new data, subsequent findings, or other methodological considerations may lead to future modifications.

### Introduction

Cognitive impairments are a core feature of schizophrenia and a major determinant of poor functional out-

come (1,2). No current pharmacological treatments expressly target these impairments. Conventional antipsychotics have only modest effects on cognition (3). In comparison to conventional antipsychotics, second generation antipsychotics (SGAs) may have additional benefits (4), but it is not known whether these benefits represent direct cognitive enhancement or indirect effects mediated through decreased extrapyramidal and associated symptoms or through decreased need for concomitant anticholinergic agents (4,5). Regardless, patients adequately treated with SGAs continue to exhibit marked cognitive impairments. In the absence of effective treatments, development of medications to treat cognitive impairments is a high public health priority.

Responding to the need for new drugs to treat cognitive impairments, the National Institute of Mental Health (NIMH) established the initiative: *Measurement and Treatment Research to Improve Cognition in Schizophrenia* (MATRICES). The initiative has four main goals, including development of novel pro-cognitive agents and facilitation of their regulatory approval. This paper summarizes the proceedings of the April 2004 FDA-NIMH-MATRICES Workshop on Clinical Trial Designs for Neurocognitive Drugs for Schizophrenia, the purpose of which was to arrive at evidence-based suggested guidelines on clinical trial design for establishing therapeutic efficacy of two general types of potential pro-cognitive agents: adjunctive/co-treatment agents and broad spectrum agents. Adjunctive/co-treatment agents have cognitive impairments as their primary indication, whereas broad spectrum agents have cognitive impairments and other symptoms of schizophrenia as co-primary indications. The lack of consensus on clinical trial design to establish therapeutic efficacy has been a formidable barrier to drug development.

### Methods

Prior to the FDA/NIMH/MATRICES workshop, a steering committee comprised of NIMH, FDA, and MATRICES scientists selected experts in these areas: cognitive impairments in schizophrenia, neurocognition, neuropharmacology, clinical trial methodology, and biostatistics. The experts were asked to address specific questions relating to clinical trial design of adjunctive/co-treatment and broad spectrum agents. Questions were

grouped under three topic areas: 1) inclusion criteria; 2) outcome measures; and 3) other design and statistical issues. At the workshop, experts reviewed relevant evidence before offering the discussion panel their proposed guidelines for a given subset of questions. The discussion panel, which consisted of the expert presenters, steering committee members, and representatives from FDA, academia, and industry (see list at end of manuscript), deliberated to reach consensus on suggested guidelines. When evidence was insufficient, suggested guidelines represent the opinion of a diverse cross-section of experts. This paper is the consensus of presenters, steering committee members, and the discussion panel.

**Results**

This section provides each question regarding clinical trial design, followed by the suggested guideline and its rationale.

*Inclusion Criteria*

*Question 1: DIAGNOSIS. Among the major psychotic disorders, what is the evidence that schizophrenia and related disorders are characterized by a unique pattern of cognitive impairments?* Suggested Guideline. Schizophrenia and schizoaffective disorder share a similar pat-

tern of cognitive impairments, which is distinct from patterns in major depression, bipolar disorder, and Alzheimer’s dementia. However, because the current approach of the FDA to approval/labeling reflects greater diagnostic specificity of the targeted population, studies of potential cognitive-enhancers should initially include only patients with schizophrenia.

Rationale. Meta-analyses suggest that patients with schizophrenia are characterized by a distinctive pattern of cognitive impairment (2,6–11; see Figure 1), which involves maximal impairment in memory, attention, and reasoning and problem solving. General verbal ability (e.g. vocabulary) and visual spatial ability are less impaired. Specialized memory retention indices do not reveal greater impairments than more immediate recall memory measures, suggesting that initial encoding difficulties rather than differential forgetting may be primarily responsible for memory impairments. Patients with schizoaffective disorder show a very similar pattern to that observed in schizophrenia (12,13).

In contrast, the characteristic profile in major depressive disorder emphasizes maximal impairment in verbal fluency, delayed episodic memory, and effortful, timed tasks (7). General verbal ability is relatively preserved, and problem-solving flexibility is minimally impaired. The pattern of cognitive impairment in patients with bipolar disorder, especially in the euthymic state, also differs

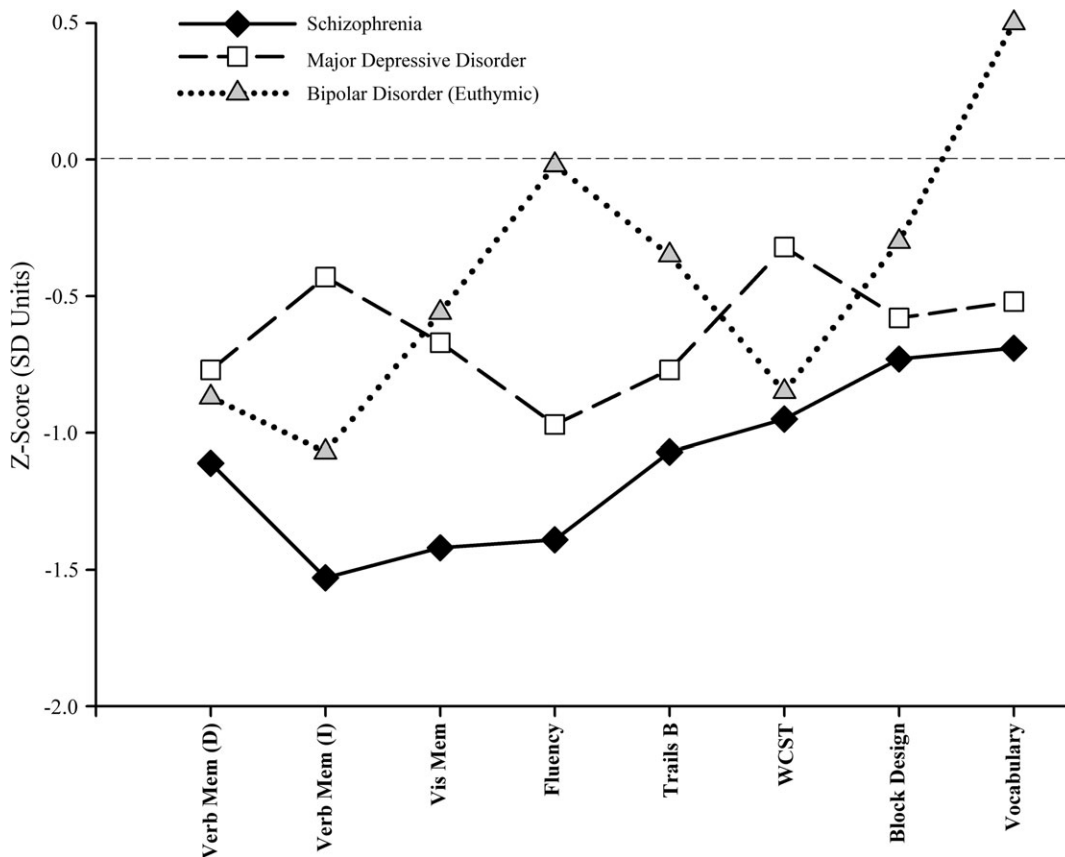


Fig. 1.

from that of schizophrenia (8). Euthymic bipolar patients show a selective cognitive impairment pattern that is distinctive from the typical broad, enduring pattern of cognitive impairments in schizophrenia (9; see Figure 1). A further distinctive feature of cognitive impairments in schizophrenia is that several core cognitive impairments in schizophrenia are relatively stable across fluctuations in clinical symptoms. This characterizes early perceptual encoding and sustained, focused attention across psychotic and fully remitted clinical states (14) and several other cognitive functions across less dramatic changes in clinical states (15–18). Thus, in contrast to mood disorders, the presence of cognitive impairments in schizophrenia is relatively independent of clinical state, though the severity of the impairment may fluctuate with change in symptoms (see below for further discussion of this issue).

Finally, meta-analysis shows that the neurocognitive profile in Alzheimer's dementia is characterized by very large deficits in visual and verbal immediate recall and distinctive additional deficits in retention of newly learned information that do not occur in schizophrenia (10). Specialized indices of memory retention show even more dramatic differences between Alzheimer's dementia and schizophrenia than do typical neuropsychological test scores (11).

*Question 2: CLINICAL STATE AND SYMPTOM SEVERITY. What design approaches should be used to isolate change in neurocognitive domains from changes in other symptom domains?*

*Suggested Guideline.* To isolate change in cognitive function from change in symptoms and other clinical features, include subjects who: a) have been clinically stable and in the residual (non-acute) phase of their illness for a specified period of time (e.g., 8–12 weeks); b) have been maintained on current antipsychotic and other concomitant psychotropic medications for a specified period of time sufficient to minimize potential complications of assessment of cognitive status (e.g., 6–8 weeks) and on current dose for a specified time period (e.g., 2–4 weeks); c) have no more than a “moderate” severity rating on hallucinations and delusions (e.g., Brief Psychiatric Rating Scale (BPRS) Hallucinatory Behavior or Unusual Thought Content item score # 4); d) have no more than a “moderate” severity rating on positive formal thought disorder (e.g., BPRS Conceptual Disorganization item score # 4); e) have no more than “moderate” severity rating on negative symptoms (e.g., all Scale for the Assessment of Negative Symptoms global items 3 or Positive and Negative Syndrome Scale–negative syndrome total score # 15); and e) have a minimal level of extrapyramidal symptoms (e.g., Simpson-Angus Scale total score # 6) and depressive symptoms (e.g., Calgary Depression Scale total score # 10). These suggested guidelines apply to studies of either adjunctive/co-treatment or broad spectrum agents.

*Rationale.* A major methodological issue in evaluating drug efficacy for cognitive impairment is the isolation of the drug effect on cognition from other concurrent changes in clinical status that also may affect cognitive performance. The failure to distinguish between a specific effect on cognition from a more general effect on clinical status may lead to erroneous conclusions. Findings from cross-sectional and longitudinal studies in patients with schizophrenia can guide specific design approaches by providing an estimate of the nature and magnitude of relationships between cognition and symptoms. The relationship is not necessarily uniform across different symptom domains, i.e., certain symptom domains may be more strongly associated with particular cognitive functions than others, or the direction of association may vary across symptom domains and cognitive function. Therefore, cross-sectional and longitudinal studies are discussed below by these symptom domains: a) hallucinations and delusions; b) disorganized behavior, which includes positive formal thought disorder, inappropriate affect, and bizarre behavior; and c) negative symptoms (19).

*Cross-sectional Studies.* In general, hallucinations and delusions are not associated with cognitive impairment (20,21). Disorganized behavior is somewhat more associated with cognitive impairment, especially measures of IQ, executive function, and language (22). In contrast, negative symptoms have a consistent and robust relationship with cognitive impairments (21–24), especially when distinguishing between primary, enduring negative or deficit symptoms and secondary negative symptoms (25,26). A fairly consistent relationship is found between deficit symptoms and neuropsychological measures of disturbed frontal and parietal lobe function (26,27), raising the possibility that the relationship could be causal, or could emerge from a common neural substrate (25).

*Longitudinal Studies.* Longitudinal studies are of three types: a) studies examining change in symptoms and cognitive function across drug-free and medicated conditions; b) clinical trials examining change in symptoms and cognitive function across different medications; and c) studies examining change in symptoms and cognitive function in the absence of an intervention.

*Drug-free versus Medicated Studies.* In contrast to cross-sectional studies, longitudinal studies that compared patients when they were drug-free to when they were medicated, found significant associations between positive psychotic symptoms and cognitive impairments (28–30, but not 31). These associations were usually in the small to moderate range, were more pronounced for disorganized behavior than for hallucinations and delusions, and involved cognitive measures of attention, memory and language (29,30).

A similar picture emerged for negative symptoms. Small to moderate significant correlations were observed between changes in negative symptoms and changes in

attention, language, processing speed, reasoning/problem-solving, spatial ability, and verbal and visual memory measures (29–31; but see 28). The Cannon and colleague study (28) found no significant negative symptom changes across medication status, which may have precluded the ability to detect such associations.

In the Weickert and colleague study (30), there was a significant 12 point difference in full-scale IQ between medicated and unmedicated patients. The IQ change was not associated with change in symptoms. The marked change in IQ, in the absence of a significant correlation with symptom change, suggests that medications may enhance the ability to take tests, since antipsychotic medications are not known to produce such dramatic changes in IQ.

**Clinical Trials.** A series of studies examined changes in symptoms and cognition in clozapine trials of partially-responsive or treatment-resistant patients with schizophrenia. In general, the degree of symptom change in these studies is less than that observed in studies that involve a drug-free period. The majority of these studies failed to document a significant relationship between change in positive symptoms and change in cognition (32–35; but see 36). In contrast, change in negative symptoms was moderately associated with change in either verbal or visual memory, processing speed, or reasoning/problem-solving measures (32,33,36). However, two studies failed to find any associations between change in negative symptom and cognitive measures (34,35).

Similarly, in treatment-resistant patients, Bilder and colleagues (37) reported that change in negative symptoms, but not positive symptoms, was significantly associated with change in working memory, reasoning/problem-solving, visual memory, processing speed. In first-episode patients, Keefe and colleagues (38) reported that changes in negative symptoms and extrapyramidal symptoms were significantly associated with change in their global neurocognitive measure in patients randomized to haloperidol but not to olanzapine. They also failed to find an association between change in depressive symptoms and change in cognition, but the study was not designed to address this.

**Longitudinal follow-up Studies.** In a 5-year follow-up study of outpatients, there were no significant changes in clinical status and no significant relationships between change in either positive or negative symptoms and change in cognitive function (39).

In summary, studies suggest that the likelihood of detecting an association between change in symptom status and cognition varies with the magnitude of change in clinical status. In the absence of clinical change, there are no associations (39), whereas in studies with marked change in clinical status, there are significant associations between positive and negative symptoms and cognitive measures (28–35,36). In addition, the Weickert and colleague study (30) suggests that there may be a general

ability to take tests, which is associated with more stable clinical states and is not adequately captured by symptom measures. Finally, few studies examined cognition in relation to change in affective symptoms or side effects.

*Question 3: ANTIPSYCHOTIC MEDICATIONS. Which antipsychotic(s) should be allowed in studies of an adjunctive/co-treatment agent?* Suggested Guideline. Select an antipsychotic(s) depending on the adjunctive/co-treatment agent and the stage of study. In stage I, avoid pharmacodynamic or pharmacokinetic interactions between the adjunctive/co-treatment agent and the antipsychotic. In stage II, evaluate the impact of potential pharmacodynamic or pharmacokinetic interactions with a larger, stratified sample, in an all-comers design, with few if any restrictions on allowed antipsychotics.

**Rationale.** Selection of an antipsychotic(s) can be based on a two-stage approach to studying potential interactions with the adjunctive/co-treatment agent. In stage I, the primary goal is proof-of-concept of the agent (after safety and pharmacokinetic properties of the adjunctive/co-treatment agent have been established). This stage should minimize confounding variables in order to maximize likelihood of detecting therapeutic effects. In stage II, the focus is on demonstrating effectiveness of the agent under typical clinical conditions, which may include identifying factors limiting effectiveness or tolerability.

Regarding pharmacokinetic interactions, SGAs are unlikely to affect metabolism of the adjunctive/co-treatment agent, since they are not known to induce or inhibit hepatic cytochrome P450 enzymes. But adjunctive/co-treatment agents that inhibit or induce hepatic metabolism could significantly affect levels of SGAs. Examples include a reported four-fold elevation of serum levels of clozapine or quetiapine following administration of agents that inhibit hepatic cytochrome P450 enzymes 1A2 or 3A4 (40,41), and 75% or greater reduction in serum concentrations of these two agents after administering anticonvulsant agents that induce cytochrome P450 enzymes (42). The conventional antipsychotics inhibit metabolism of other drugs, such as tricyclic antidepressants, possibly by effects on cytochrome P450 2D6 (43).

Pharmacodynamic interactions are more likely than pharmacokinetic interactions to complicate interpretation of drug effects and dose-response relationships. SGAs bind with varying affinities to a broad range of receptors, including several receptors that are potential targets for cognitive-enhancing agents (44–47). Combining a putative cognitive-enhancing agent with an antipsychotic with relatively high affinity for the targeted receptor could attenuate potential therapeutic effects. Agonists at the strychnine insensitive glycine receptor, such as glycine and D-cycloserine, illustrate such pharmacodynamic interactions; these putative cognitive-enhancing agents produce significant improvement in

negative symptoms when added to most antipsychotics, but fail to produce an effect or even worsen negative symptoms when added to clozapine (48–51). Similarly, olanzapine and clozapine may compete with drugs acting at nicotinic acetylcholine receptors as a result of potent release of acetylcholine (52).

In stage II studies, the generalizability of the efficacy and tolerability of the cognitive-enhancing agent may be evaluated through the study of representative samples of patients under typical clinical conditions. The potential impact of combining the investigational drug with an antipsychotic that has a similar mechanism of action or a mechanism of action that could potentially undermine the mechanism of action of the investigational drug should be assessed, ideally in a stratified sample that allows examination of specific combinations.

*Question 4: POLYPHARMACY. Should antipsychotic polypharmacy be allowed in studies of an adjunctive co-treatment agent?* Suggested Guideline. Exclude subjects who are taking more than one antipsychotic. Antipsychotic combinations are an unnecessary complication in studying an adjunctive/co-treatment agent.

Rationale. Combining cognitive-enhancing agents with multiple antipsychotics is an unnecessary complication. Potential interactions between investigational agents and individual antipsychotics can be studied in a large, stratified sample of patients treated with antipsychotic monotherapy, whereas inclusion of patients treated with antipsychotic polypharmacy may reduce the study's power to examine potential drug interactions. Polypharmacy is relatively common in the United States where surveys have found two or more antipsychotics prescribed to about 15% of outpatients with schizophrenia and to as many as 50% of inpatients (53).

*Question 5: CONCOMITANT MEDICATIONS. Should concomitant medications be allowed?* Suggested Guideline. In stage I, avoid pharmacokinetic and pharmacodynamic interactions between the adjunctive/co-treatment agent and any concomitant medications (e.g., SSRIs). In stage II, examine potential pharmacokinetic and pharmacodynamic interactions on an agent-specific basis.

Rationale. In stage I studies, the same principles of study design apply to concomitant medications as with antipsychotic drugs (see Question 3). Potential pharmacokinetic interactions should be considered with the selective serotonin reuptake inhibitors and with anticonvulsant medications (42,54). Concomitant medications that potentially interact pharmacodynamically with the adjunctive/co-treatment agent include: serotonergic or noradrenergic antidepressants; muscarinic anticholinergic agents; glutamatergic NMDA receptor antagonists; dopaminergic or noradrenergic psychostimulants; and benzodiazepines and anticonvulsants acting

on GABAergic and glutamatergic systems. Excluding all concurrent medications in stage I trials may not be feasible, since it would exclude the large majority of potential subjects in most schizophrenia patient samples.

*Question 6: MAXIMUM LEVEL OF IMPAIRMENT. In order to detect a therapeutic effect, should a maximum level of cognitive impairment be specified?* Suggested Guideline. Exclude patients from a trial only if their cognitive impairment severity compromises the validity of the cognitive outcome measures. This determination may be made using a combination of clinical judgment and objective data, such as premorbid IQ scores and premorbid IQ estimates from current reading level.

Rationale. The rationale for this exclusion criterion is based on two considerations: some patients may be so severely cognitively impaired that: a) they cannot benefit from cognitive enhancement; and b) their data may not be valid.

There is equivocal evidence on the relationship between baseline severity of cognitive impairments and treatment-related cognitive improvement. Recent data suggest that patients with the least severe cognitive impairments may demonstrate the most benefit from antipsychotic treatment (e.g., 55), while data from other studies have suggested that patients with the most severe cognitive impairments may benefit most (e.g., 38,56). Thus, the empirical evidence is not strong enough to exclude *a priori* all patients with severe cognitive impairments.

In contrast, there are clearly patients whose cognitive impairments are so severe that they cannot be assessed validly on neuropsychological tests. The assessment of their neuropsychological test validity should be based on the clinical judgment of the neurocognitive tester, psychologist, and/or treating psychiatrist. For instance, severely impaired patients judged incapable of understanding the test instructions are not testable. Patients with evidence of Pervasive Developmental Disorder, such as a premorbid IQ score below 70 or a documented diagnosis, should be excluded because they may have cognitive impairments unresponsive to a schizophrenia-directed treatment. Premorbid IQ estimates based upon measures of current crystallized intelligence, such as information or reading scores, may deflate an IQ estimate and should only be used with great caution as screening criteria. Patients with poor reading scores may not be able to understand how to perform some neurocognitive tests, calling into question the validity of test results.

*Question 7: MINIMUM LEVEL OF IMPAIRMENT. In order to detect a therapeutic effect, should a minimum level of cognitive impairment be specified in the inclusion criteria?* Suggested Guideline. Exclude subjects from a trial if their level of cognitive functioning is so high

that they perform at or near ceiling and therefore cannot demonstrate improvement. With a properly constructed cognitive test battery, this level of performance will be very rare.

**Rationale.** If a patient performs perfectly or near perfectly on a test battery (i.e., the “ceiling effect”), then he or she is not a good candidate for inclusion, since cognitive enhancement cannot be detectable. A cutoff of one standard deviation below perfect performance on a test or test battery will help to exclude patients unlikely to demonstrate improvement. However, in a carefully designed battery, this level of performance almost never occurs.

There is insufficient evidence to exclude patients who perform well, but not at or near ceiling. The few relevant cognitive enhancement trials report mixed results on the relationship between the level of cognitive impairment and response to treatment (55–57). There is also no reason to exclude patients defined as “unimpaired” relative to healthy controls (e.g. performance that is 1.5 standard deviations below the healthy control mean), since almost all of these patients have cognitive impairments relative to what would be expected had they never developed schizophrenia (58). Defining cognitive impairment as a failure to meet one’s expected level of cognitive functioning means that almost all patients with schizophrenia (98%) qualify as cognitively impaired (Keefe *et al.*, unpublished data, 2004). Further support for this notion comes from the observation that almost all affected monozygotic twins perform worse on cognitive tests than their unaffected co-twins (59). Thus, almost all patients with schizophrenia are likely to benefit from cognitive enhancement and should therefore be included in trials.

*Question 8: SCREENING ASSESSMENTS. If patients are screened for inclusion based upon their level of cognitive impairment, how should screening assessments be conducted?* Suggested Guideline. If a screening assessment must be used, then use an assessment that is different from the measure used to assess cognitive outcome during the trial. The primary exception would be if a multiple baseline strategy is employed.

**Rationale.** The use of screening instruments for inclusion must address potential practice effects (i.e., improved test performance due to repeated exposures to the test), novelty effects (i.e., decreased test performance due to lack of familiarity with the test demands, which lessens with repeated exposures to the test), and the natural tendency of deviant performances to regress toward the mean upon further testing (i.e., “regression to the mean”). Practice and novelty effects are particularly important in situations in which change over time is a key outcome measure, as is the case with most cognitive enhancement trials in schizophrenia. “Regression to the mean” effects may potentially lower the sensitivity of

the screening instrument. Because error variance for any cognitive measure creates deviation from the “true score” of a patient, a certain percentage of patients who meet a minimum level of cognitive impairment criterion will, by chance factors alone, regress toward their true score when testing is repeated. However, a similar percentage of patients who miss the inclusion criterion will not have the opportunity to be tested again, and thus will not counteract the effects of the included group. These tendencies could cause a study-wide inflation of test performance, and may potentially reduce the sensitivity of the cognitive outcome measure.

If, despite concerns about using a screening assessment, a decision is made to screen patients, then the optimal strategy would be to use a screening instrument other than the primary outcome measure. While the effects of practice, novelty, and regression to the mean will be lessened if a different neurocognitive instrument is used for screening, these effects are general, and may still occur. A very brief reading assessment may provide information about which patients cannot be tested or who have such a low level of intellectual function that they do not understand the test.

There are three other potential screening strategies. The first strategy uses the outcome neurocognitive battery as the screening instrument. This approach will lessen novelty effects, since repeated baseline assessments will help reduce variance in cognition due to unfamiliarity with the test-taking process. However, the sensitivity of the battery may be reduced, because the effects of regression to the mean and practice effects will be at their maximum. The second strategy uses a subset of the primary outcome measure for screening. This approach will lessen the impact of the test-specific component of factors, such as regression to the mean, and reduce the duration of the screening process. However, it could produce different practice and novelty effects across tests and test domains, which could reduce the ability to differentiate treatment effects across different cognitive domains. Finally, the use of multiple assessments at baseline would minimize novelty effects and maximize practice effects. As a result, any improvement in cognitive function would be more clearly attributable to the pharmacological intervention. However, this method would be used not to screen patients, but to eliminate error variance.

### *Outcome Measures*

In the absence of a standard neuropsychological test battery for use in clinical trials, NIMH convened academic, industry and regulatory experts under the MATRICS initiative to review available data and create a consensus battery specifically designed to assess cognitive impairments in schizophrenia (60). The MATRICS cognitive battery will be used to assess the primary

outcome measure: change in cognitive performance. The battery will assess the following seven domains: attention/vigilance, reasoning and problem solving, speed of processing, social cognition, verbal learning and memory, visual learning and memory, and working memory. The MATRICS Psychometric and Standardization Study (PASS) study is currently evaluating the psychometric properties of the beta version of the battery. PASS results will be used to select 1–2 tests in each cognitive domain for the final MATRICS consensus cognitive battery. This process will be completed by fall of 2004. This battery will be about 90 minutes in length and is expected to be administered in its entirety.

**Question 9: CO-PRIMARY OUTCOME MEASURES.**

*What are the issues regarding inclusion of a co-primary outcome measure of community functioning?* Suggested Guideline. If FDA requires a co-primary measure, the measure should assess a clinically meaningful/relevant functional outcome, but not necessarily community functioning. Community functioning is highly dependent on psychosocial services, patient skills, and social support – factors that are usually beyond the control of clinical trials.

Rationale. The current position of the FDA is that concurrent change on a co-primary measure of functional outcome will be required for approval of a neurocognitive drug for schizophrenia. The FDA may be willing to accept a co-primary measure with good face validity, whether a proxy measure of community outcome or an interview-based measure of cognition, before formal validation of a co-primary measure is completed.

The argument in favor of a co-primary outcome measure is the importance of a face valid measure of functional improvement. Clinicians and consumers may not be able to appreciate changes on cognitive performance measures alone. To increase acceptance of the drug, it may be necessary to demonstrate effects on measures that reflect clinically meaningful improvement. In addition, improved functional outcome is the ultimate goal of cognitive-enhancing drugs. Finally, although co-primary measures are not routinely required for drug development programs, there is a precedent for requiring functional or global outcomes in certain situations. The Division of Neuropharmacological Drug Products (DNBP) at FDA has required such outcomes in drug development programs for the treatment of cognitive impairment in Alzheimer’s disease and for the treatment of psychosis of Alzheimer’s disease (61).

The arguments against a co-primary measure of functional outcome, including community (e.g., work and social) outcome, relate to psychometric and conceptual concerns. First, there is a measurement problem: the reliability and/or validity of self-report measures of functional status are not well established. Second, there is increasing evidence that mediating variables (e.g., coping

ability, skill acquisition, social cognition) act between measures of cognition and functional outcome (62–64). These variables may obscure the translation of cognitive-enhancing effects into changes in functional status. Third, changes in community functioning are far removed from the biological systems that would be altered by cognitive-enhancing drugs. They heavily depend on factors that are typically beyond the control of clinical trial studies, including availability of psychosocial rehabilitation, social support networks, community opportunities (e.g., local employment rates) and educational opportunities.

The other major issues associated with the use of co-primary measures are the statistical implications. The separate assessment of cognition and functional outcome will need to be addressed in statistical analysis plans. There are two basic decision rules that pre-specify whether eventual results qualify as positive: a) the “Or Rule,” which requires superiority on either co-primary outcome; and b) the “And Rule,” which requires superiority on both co-primary measures. The “Or Rule” requires an alpha adjustment to guard against elevated risk of Type I error, or false positive results. A consequence of the alpha adjustment is a reduction in statistical power. However, such reductions can be prevented if projected sample size requirements are based on the multiplicity-adjustment (65). But multiplicity adjustment increases sample size requirements and corresponding costs, both monetary and number of subjects exposed to risks of a randomized clinical trial. The “And Rule” is the current standard of the DNBP. This rule needs no explicit alpha adjustment. If the “And rule” is applicable, then the investigator must be careful to select well-validated scales that separate in pilot studies, since both co-primary measures based on these instruments must be positive.

**Question 10: CO-PRIMARY MEASURE CHARACTERISTICS.**

*What are the ideal characteristics of a co-primary outcome measure of functional improvement?* Suggested Guideline. A co-primary measure should have the following characteristics: a) good face validity for patient improvement; b) expected to change in close temporal proximity to changes on cognitive performance measures; c) not be heavily dependent on range of rehabilitation opportunities and level of social support; and d) practical (from the perspective of the experimenter) and tolerable (from the perspective of the subject).

Rationale. Although community outcome may be too far causally removed from the actions of cognitive-enhancing drugs, other options may be appropriate. There may be measures that reflect real-time changes in underlying neural systems and do not depend on levels of psychosocial support and community opportunities. Two potential types of co-primary measures are: 1) assessment

of functional capacity; and 2) interview-based assessment of cognition.

Functional capacity assessments may use props and role playing to assess whether a person can maintain a social conversation, prepare a meal, take public transportation, or manage their medications. These assessments are simulated activities conducted in the clinic and do not rely on observing the individual in the community (66,67). If patients perform well on a measure of functional capacity (also called “proxy” measures of outcome), that does not guarantee they will be able to perform the tasks in the community. It only means that they could perform the task in the community. Changes in functional capacity are likely to occur more closely in time with changes in underlying cognitive performance.

Interview-based measures of cognition are structured or semi-structured interviews that could be administered to patients, care-givers, or clinicians. Patients are asked about their subjective view of their cognitive abilities, or their difficulty in performing tasks of everyday living (e.g., reading a book, remembering where objects are placed, etc.). Informants also can be interviewed to rate their impressions of subjects’ cognitive abilities.

*Question 11: VALIDITY OF PROXY MEASURES. What is the best approach to assess the validity of proxy measures of functional outcome and interview-based measures of cognition?* Suggested Guideline. Potential co-primary proxy measures have: a) good test-retest reliability; b) demonstrated associations with cognitive performance measures; and c) demonstrated associations with community functional status.

Rationale. Test-retest reliability is the most important property of a test used in randomized clinical trials (60). If a test has poor test-retest reliability, it is difficult to demonstrate treatment effects, and its validity cannot be determined. In addition, a proxy measure of functioning should be related to cognitive performance measures, because it would be expected to change in real-time with underlying changes in cognitive abilities. Proxy measures should also be related to assessments of community functioning, because they are intended to be more proximal indications of how well patients function in their daily lives.

How well do potential proxy measures (i.e., functional capacity and interview-based measures of cognition) meet these criteria? Unfortunately, we do not yet know. In schizophrenia and other disorders, several studies have found significant cross-sectional associations between cognitive performance and proxy measures of social problem-solving and daily activities (67,68–71). However, there are few data to link functional capacity measures to community outcome. While there are a large number of studies on interview-based measures of cognition in schizophrenia and neurological disorders, many

studies fail to find associations between interview-based measures and cognitive performance, or they find relationships between the interview-based measures and dysphoric mood, instead of cognitive performance (72–76).

In summary, the field does not have sufficient validity data on any potential co-primary measure of functional outcome, either community function, functional capacity, or interview-based measure of cognition. It is possible that decisions concerning use of co-primary measures will occur as part of a longer-term iterative process. Validation of co-primary measures may first require identification of drugs with potent effects on cognitive performance measures. Once such drugs are identified, they could be used to effectively establish the validity of potential co-primary measures.

#### *Other Design and Statistical Issues*

*Question 12: CHOICE OF PRIMARY MEASURE. What are issues regarding the choice of primary cognitive outcome measure(s)?* Suggested Guideline. Pre-specify a single reliable and valid primary cognitive outcome measure, either global or domain-specific, based on its psychometric properties and results of pilot studies.

Rationale. The choice of primary cognitive efficacy measure(s) should be based on the following design and statistical considerations: first, pilot studies must demonstrate that the measure(s) is(are) differentially sensitive to investigational and comparator drugs. Second, the measure(s) must be developed, validated, and psychometric results replicated before its use in a randomized controlled clinical trial (RCT). Psychometric studies must provide empirical support for the construct validity and reliability of the measure (e.g. test-retest, interrater, and/or internal consistency). The relevance of previously reported psychometric properties can be weighed by examining the design and implementation of those studies, e.g. study duration, assessment frequency, and subject status. Third, because multisite RCTs typically have at least as many raters as sites, interrater reliability must be established prospectively during the pilot phase of each RCT. Fourth, tradeoffs between global cognitive measures and individual cognitive domain measures must be weighed. One rationale for using a global measure is that its internal consistency reliability tends to increase with multiple correlated items. As the measure’s reliability increases, within-group variability decreases. Consequently, the between-group effect size increases and sample size requirements decrease (77). However, global measures have limitations: a) a multi-domain measure or battery may not be clinically meaningful; b) a global score may wash out the effect of one domain and miss a true effect; and c) unweighted totals arbitrarily place greater emphasis on some scales/items; whereas,



sample-specific weightings of scales limit both the inter-pretability and generalizability of psychometric results. Finally, the choice of a global versus specific cognitive domain measure must be specified *a priori*.

The decision to choose a composite cognitive measure or a specific cognitive domain measure will primarily depend on pilot study results. Although a composite measure will likely be optimal for most drugs, certain drugs may be better suited to a specific cognitive domain when, for example, pilot studies suggest no effect on a composite measure, yet a clear effect on a particular cognitive domain. In this case, it might be reasonable to select the particular domain as the primary cognitive measure. Alternatively, a drug may affect a composite measure, but have a more striking effect on a particular cognitive domain. In this case, it may be reasonable to select the composite measure as the primary cognitive measure, and also to select the particular cognitive domain as a key secondary endpoint. To support this secondary endpoint, additional work would be needed to establish the superiority of this drug to other cognitive-enhancing drugs for this particular cognitive domain.

*Question 13: TESTING OCCASIONS. What are the issues regarding the number of testing occasions?* Suggested Guideline. Use more rather than fewer testing occasions to reduce the impact of attrition and to capture change in symptom severity over time. Use data analytic procedures that incorporate repeated measures of the outcome.

Rationale. The number of testing occasions is partly a function of trial duration. Longer trials need more assessments to capture change in cognitive function over time and to mitigate the impact of attrition. However, unless the data analytic procedure incorporates repeated measures, the alpha level must be adjusted for the multiple statistical tests corresponding to each of the repeated assessments over the course of the trial. Unlike symptom severity ratings, cognitive tests cannot be administered over numerous consecutive weeks and therefore, one must consider the number of available parallel forms. The frequency of cognitive assessments is limited by potential practice and novelty effects. The influence of practice effects, in part a function of the time between test administrations, is a concern, which is offset to some extent by the use of a comparator treatment arm. In addition, a novelty effect could enhance or impair performance during the initial administration of the assessment.

*Question 14: HETEROGENEITY OF SEVERITY AND RESPONSE. In light of the expected heterogeneity of severity and response within and across different cognitive domains, what approaches to design and analysis should be used to detect a therapeutic effect?* Suggested Guideline. In order to reduce baseline within-group het-

erogeneity and to increase the chance of detecting a therapeutic effect, include subjects in the residual (non-acute) phase of their illness and use one primary efficacy measure.

Rationale. Establishing minimum and maximum symptom severity inclusion criteria reduces the baseline within-group heterogeneity, likely decreases the within-group heterogeneity at end-of-study, and, thus, increases the chance of detecting a therapeutic effect. The problem of heterogeneity of response, on the other hand, stems from concern about inconsistent results from having multiple primary outcome measures. As stated earlier, in accord with the “And Rule,” if co-primary outcome measures are delineated in the protocol and the investigational drug is not superior on each of those measures, the trial will be deemed negative. The problem of inconsistency is readily circumvented by identifying just one primary efficacy measure.

*Question 15: CONCURRENT CHANGE IN SYMPTOMS. What approaches to design and analysis should be used to control for potentially concurrent changes in other symptom domains?* Suggested Guideline. Statistical approaches cannot be used to rule out pseudospecificity (i.e., an artificially narrow claim of cognitive enhancement that could result from post-baseline confounding, such as reduction in other aspects of the illness; 78). Pseudospecificity is best dealt with by restricting symptom severity prior to randomization.

Rationale. Several statistical approaches attempt to disentangle concurrent changes across different outcome domains. However, there are limitations associated with attempting to control for post-baseline confounding variables. We will review several statistical approaches, which attempt to deal with the problem of concurrent symptom change and subsequent concerns of pseudospecificity.

A common statistical approach is to use covariate adjustment to control for symptom changes. The particular method of covariate adjustment will vary, based on the form of the dependent variable, and include analysis of covariance, multiple linear regression analysis, logistic regression, and Cox’s proportional hazards models. Regardless, the International Conference on Harmonization warns that the use of post-baseline covariates is “not advisable” and should not be used to rule-out pseudospecificity (79). Among other reasons, this position stems from the problem of inferring causality through statistical adjustment for post-randomization covariates.

Path analysis is a correlational approach that involves a series of regression equations, with one equation for each dependent variable in the path model (80). In essence, it is simply another strategy of covariate adjustment and cannot be used to infer causality in the presence of a confounding variable in a randomized experiment and, thus, cannot be used to rule out pseudospecificity.

Furthermore, the causal direction between two domains assessed contemporaneously is ambiguous.

A simple, yet effective strategy to remove the effect of a confounding variable on treatment effect estimate is to stratify group assignment by pre-specified levels of the variable (81). In the case of symptoms, this would involve separate efficacy analyses for those with mild, moderate, and severe baseline symptoms. If there is not an interaction between treatment and baseline severity, the results can be pooled using a procedure, such as the Mantel-Haenszel approach (see 82). Unfortunately, stratification on baseline severity will not provide evidence of specific cognitive benefit if there is post-randomization reduction in severity of symptoms.

*Question 16: COMPARISON GROUP. What are the appropriate comparator agents for adjunctive/co-treatment agents and broad spectrum agents?* Suggested Guideline. To study an adjunctive/co-treatment agent, use placebo as the comparator. The choice of comparator for a broad spectrum agent poses a more substantial challenge, but should be, at worst, cognitively neutral.

Rationale. The choice of comparator differs between adjunctive/co-treatment agents and broad spectrum agents. The selection of a comparator for an adjunctive/co-treatment agent is relatively straightforward: one group of subjects is randomized to receive the combination of antipsychotic and adjunctive/co-treatment agent, whereas the other group receives antipsychotic and adjunctive placebo. In this design, the antipsychotic is a constant and the adjunctive treatment varies, i.e., active agent or placebo. Superiority of the adjunctive/co-treatment agent can be interpreted unambiguously.

The selection of comparator for a broad spectrum agent, which has both cognitive impairments and other symptoms of schizophrenia as co-primary indications, is more complicated. The evaluation of the cognitive-enhancing effects of a broad spectrum agent should be separated from the evaluation of its antipsychotic efficacy. The assessment of antipsychotic efficacy should follow standard designs for antipsychotic agents in acutely ill patients. The assessment of cognitive-enhancing effects should utilize stabilized patients in order to minimize the secondary effects of acute stabilization of psychosis and agitation on cognition. The ideal comparator agent should be, at worst, cognitively-neutral. The comparator could even have cognitive-enhancing effects, but such a drug would require a larger sample size to demonstrate a significant difference between it and the broad spectrum agent.

There are three potential comparator agent options for studying a broad spectrum agent: a) placebo; b) conventional antipsychotic; and c) SGA. The use of placebo minimizes potential confounding from concurrent decrease in symptoms or neurological side effects. However, placebo use is compromised by the potential for signifi-

cant symptom exacerbation, especially in long duration trials. The use of conventional antipsychotics raises questions about interpretation of study results, because extrapyramidal and other neurological side effects may act to impair cognition performance (37,38,83). In addition, conventional antipsychotics are commonly used in conjunction with anticholinergic agents to minimize extrapyramidal side effects. The use of anticholinergic agents may further complicate clinical trials, because these agents can adversely affect memory (84). Further, the broad spectrum agent might provide more pronounced symptom relief, resulting in less cognitive impairment. The use of a SGA minimizes potential confounding due to neurological side effects, but still poses difficulties. Any apparent cognitive improvement found with the broad spectrum agent relative to a conventional antipsychotic or SGA is difficult to interpret in the absence of a placebo control. One interpretation could be that the broad spectrum agent improves cognition while the comparator either does not, or does not improve cognition as much. A second interpretation could be that the broad spectrum agent has no effect on cognition, but the comparator actually impairs cognition. A third interpretation could be that both agents impair cognition, but the broad spectrum agent is less impairing. In light of the potential difficulty of interpretation, any finding of cognitive enhancement with a broad spectrum agent may not be appropriate for the *Indications and Use* section of the drug's label, because it would imply that cognitive enhancement has been established. It may be more appropriate for the *Adverse Reactions* section of the label, noting that, although the absolute effect of the broad spectrum agent on cognition cannot be determined, it appears to have significantly less cognitive liability than certain comparator agents.

*Question 17: TRIAL DURATION. What is the optimal clinical trial duration of an agent targeting cognition in schizophrenia?* Suggested Guideline. The trial needs to be of sufficient duration to show an enduring effect on cognition (i.e., at least 6 months). Longer duration studies should use multiple testing occasions, which require the existence of parallel forms of the outcome measure. Statistical procedures should be used that incorporate data from multiple assessment times.

Rationale. The FDA DNDP has stated that an enduring effect on cognition (i.e., at least 6 months) must be documented in order to apply for an indication for cognitive enhancement in schizophrenia. The primary outcome measure must be able to be administered multiple times over the course of the trial, either through the use of one form or alternate parallel forms of the measure. In addition, the statistical procedures described in the protocol must be capable of accommodating data from multiple assessment times. For instance, one of

the mixed-effects models may be selected based on the form of the dependent variable (85–89). This general strategy provides more statistical power, or reduces sample size requirements, compared with strategies that simply examine baseline and endpoint assessments (90).

Another advantage of this general data analytic strategy is the flexibility of the models to include subjects with varying numbers of post-baseline assessments (e.g., monthly assessments). This is quite useful in longer trials, which are more vulnerable to patient attrition. However, mixed-effects models only provide valid inferences with ignorable dropout (91). This refers to dropout that is a function of either previous values of the dependent variable or observed covariates (92).

Regardless of data analytic strategy, the protocol should specify that assessments should continue for the entire course of the RCT, regardless of protocol adherence (93). If a subject drops out in the second month, efforts should be made to continue assessments for entire trial 6-month duration. Unlike RCTs with patients in the acute phase, this approach could be feasible for RCTs with patients in the residual phase. This strategy serves to adhere better to the intention-to-treat principle than last observation carried forward analyses.

## Discussion

The FDA/NIMH/MATRICES workshop was designed to produce a set of study design guidelines that facilitate the development of new, innovative treatments for cognitive impairments in patients with schizophrenia (see: Innovation/Stagnation: Challenge and opportunity on the Critical Path to New Medical Products. <http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html>). The workshop produced a series of suggested guidelines for inclusion criteria, outcome measures, and design and statistical approaches. There was considerable consensus regarding diagnostic and concomitant medication inclusion criteria and on the use of cognitive screening measures. There was also a convergence of evidence to support limiting inclusion to patients in the residual phase of their illness and who present with a predetermined level of positive, negative, and depressive, and extrapyramidal symptoms. The results of studies examining the relationship between change in symptoms and change in cognition and considerations of different statistical approaches to data analysis strongly argued for this suggested guideline. A major implication is that separate studies will need to be conducted to demonstrate symptom efficacy and cognitive-enhancing effects for broad spectrum agents.

There was considerably less agreement about the feasibility of including a co-primary measure of community outcome. The importance of a face valid measure of functional improvement is counterbalanced by the lack of

a validated measure of community outcome for use in clinical trials. The FDA may be willing to accept co-primary measures that have good face validity, whether they are proxy measures of functional outcome or interview-based measures of cognition, even before the formal validation process has been completed. In addition, statistical considerations, including the requirement to demonstrate efficacy for both co-primary measures, raise serious concerns about the use of co-primary measures. The resolution of this issue will ultimately depend on the development of valid proxy measures of community outcome that are directly associated with both change in cognitive function and change in social and occupational function.

The other major unresolved issue is the choice of a comparator agent for use in studies of broad spectrum agents. The ideal choice would be the use of a cognitively-neutral agent. However, there are currently no known antipsychotic agents that meet this criterion. SGAs have less neurological side effects than conventional antipsychotics, especially high-dose conventional antipsychotic treatment, but the absence of potential adverse cognitive effects of these agents has not been definitely demonstrated. In the absence of a known cognitively-neutral agent, the unequivocal interpretation of study results is complicated.

In summary, the suggested guidelines are intended to represent a reasonable starting point for trial design of cognitive-enhancing drugs, with the understanding that there may be deviations in any particular trial, but that any deviations will need to be explained by either drawing upon new data, subsequent findings, or other methodological considerations.

## Discussion Panel

Alan Breier, M.D., Chief Medical Officer & Vice President Medical, Eli Lilly and Company; William T. Carpenter, Jr., M.D., Professor, University of Maryland School of Medicine, Maryland Psychiatric Research Center; Daniel Casey, M.D., Professor of Psychiatry and Neurology, Oregon Health & Science University; Susan Essock, Ph.D., Professor, Mount Sinai School of Medicine; Robert Freedman, M.D., Professor, Department of Psychiatry, University of Colorado Health Sciences Center; Georges Gharabawi, M.D., Group Director, Janssen Pharmaceuticals, Incorporated; John Kane, M.D., Chairman, Department of Psychiatry, Zucker Hillside Hospital; Robert O'Neill, Ph.D., Director, Office of Biostatistics, Department of Health and Human Services, Food & Drug Administration; Steve Romano, M.D., Senior Medical Director Pfizer, Incorporated; Carol Tamminga, M.D., Professor of Psychiatry, University of Texas Southwestern Medical Center

## Acknowledgements

Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) – FDA-NIMH-MATRICS Workshop on Clinical Trial Designs for Neurocognitive Drugs for Schizophrenia – Support for this conference came from NIMH Contract MH22006 (S.R. Marder, P.I., M.F. Green, Co-P.I.).

## The Authors

Robert W. Buchanan, Chief of the Outpatient Research Program at the Maryland Psychiatric Research Center (MPRC). Miriam Davis, Department of Epidemiology and Biostatistics, School of Public Health and Health Services, George Washington University. Donald Goff, Psychiatry Department, Massachusetts General Hospital, Harvard Medical School, Boston, Michael F. Green and Steve R. Marder, Department of Psychiatry and Behavioral Sciences, David Geffen School of Medicine, University of California at Los Angeles (UCLA), and VA Greater Los Angeles Healthcare System, VISN 22 Mental Illness Research, Education, Clinical Center. Richard S. E. Keefe, Duke University Medical Center. Andrew C. Leon, Weill Medical College of Cornell University, Departments of Psychiatry and Public Health, New York, NY. Keith H. Nuechterlein, Department of Psychiatry and Behavioral Sciences, David Geffen School of Medicine, University of California at Los Angeles (UCLA). The contributions of Thomas Laughren and Robert Levin were made in their private capacity; no official support or endorsement by the Food and Drug Administration is intended or should be inferred. Ellen Stover and Wayne Fenton, National Institute of Mental Health, National Institute of Health.

## References

1. Green, M.F. What are the functional consequences of neurocognitive deficits in schizophrenia? *Am J of Psychiatry* 1996; 153:321–330.
2. Heinrichs, R.W., and Zakzanis, K.K. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology* 1998 Jul; 12(3):426–45.
3. Mishara, A.L., and Goldberg, T.E. A meta-analysis and critical review of the effects of conventional neuroleptic treatment on cognition in schizophrenia: opening a closed book. *Biol Psychiatry* 2004; 55:1013–22.
4. Harvey, P.D., and Keefe, R.S.E. Studies of cognitive change in patients with schizophrenia following novel antipsychotic treatment. *Am J Psychiatry* 2001; 158:176–184.
5. Carpenter, W.T., and Gold, J.M. Another view of therapy for cognition in schizophrenia. *Biol Psychiatry* 2002; 51:969–71. Comment on: *Biol Psychiatry*. 2002; 51:972–8.
6. Nuechterlein, K.H., Barch, D.M., Gold, J.M., Goldberg, T.E., Green, M.F., and Heaton, R.K. Identification of separable

cognitive factors in schizophrenia. *Schizophrenia Research* in press.

7. Zakzanis, K.K., Leach, L., and Kaplan, E. On the nature and pattern of neurocognitive function in major depressive disorders. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology* 1998; 11:111–119.
8. Bearden, C.E., Hoffman, K.M., and Cannon, T.D. The neuropsychology and neuroanatomy of bipolar affective disorder: A critical review. *Bipolar Disorders* 2001; 3:106–150.
9. van Gorp, W.G., Altshuler, L., Theberge, D.C., Wilkins, J., and Dixon, W. Cognitive impairment in euthymic bipolar patients with and without prior alcohol dependence. *Archives of General Psychiatry* 1998; 55:41–46.
10. Zakzanis, K.K. Quantitative evidence for neuroanatomic and neuropsychological markers in dementia of the Alzheimer's type. *Journal of Clinical and Experimental Neuropsychology* 1998; 20:259–269.
11. Heaton, R., Paulsen, J.S., McAdams, L.A., Kuck, J., Zisook, S., Braff, D., Harris, M.J., and Jeste, D.V. Neuropsychological deficits in schizophrenics: Relationship to age, chronicity, and dementia. *Archives of General Psychiatry* 1994; 51:469–476.
12. Miller, L.S., Swanson-Green, T., Moses, J.A. Jr, and Faustman, W.O. Comparison of cognitive performance in RDC-diagnosed schizoaffective and schizophrenic patients with the Luria-Nebraska Neuropsychological Battery. *Journal of Psychiatric Research* 1996; 30:277–282.
13. Evans, J.D., Heaton, R.K., Paulsen, J.S., McAdams, L.A., Heaton, S.C., and Jeste, D.V. Schizoaffective disorder: A form of schizophrenia or affective disorder? *Journal of Clinical Psychiatry* 1999; 60:874–882.
14. Nuechterlein, K.H., Dawson, M.E., Gitlin, M., Ventura, J., Goldstein, M.J., Snyder, K.S., Yee, C.M., and Mintz, J. Developmental processes in schizophrenic disorders: Longitudinal studies of vulnerability and stress. *Schizophrenia Bulletin* 1992; 18(3): pp. 387–425.
15. Hoff, A.L., Riordan, H., O'Donnell, D.W., Morris, L., and DeLisi, L.E. Neuropsychological functioning of first-episode schizophreniform patients. *American Journal of Psychiatry* 1992; 149:898–903.
16. Nopoulos, P., Flashman, L., Flaum, M., Arndt, S., and Andreasen, N. Stability of cognitive functioning early in the course of schizophrenia. *Schizophrenia Research* 1994; 14:29–37.
17. Censits, D.M., Ragland, J.D., Gur, R.C., and Gur, R.E. Neuropsychological evidence supporting a neurodevelopmental model of schizophrenia: A longitudinal study. *Schizophrenia Research* 1997; 24:289–98.
18. Finkelstein, J.R.J., Cannon, T.D., Gur, R.E., Gur, R.C., and Moberg, P. Attentional dysfunctions in neuroleptic-naive and neuroleptic-withdrawn schizophrenic patients and their siblings. *Journal of Abnormal Psychology* 1997; 106:203–12.
19. Buchanan, R.W., and Carpenter, W.T. Domains of psychopathology: an approach for the reduction of heterogeneity in schizophrenia. *J Nerv Ment Dis* 1994; 182:193–204.
20. Goldberg, T.E., and Gold, J.M. Neurocognitive functioning in patients with schizophrenia, in *Psychopharmacology: The Fourth Generation of Progress*. In: Bloom, F.E., and Kupfer, D.J., eds. New York: Raven Press, Ltd 1995; pp. 1245–1257.
21. Mohamed, S., Paulsen, J.S., O'Leary, D., Arndt, S., and Andreasen, N. Generalized cognitive deficits in schizophrenia. *Arch Gen Psychiatry* 1999; 56:749–754.

22. O'Leary, D.S., Flaum, M., Kesler, M.L., Flashman, L.A., Arndt, S., and Andreasen, N.C. Cognitive correlates of the negative, disorganized, and psychotic symptom dimensions of schizophrenia. *J Neuropsychiatry Clin Neurosci* 2000; 12:4–15.
23. Addington, J., Addington, D., and Maticka-Tyndale, E. Cognitive functioning and positive and negative symptoms in schizophrenia. *Schizophr Res* 1991; 5:123–134.
24. Heydebrand, G., Weiser, M., Rabinowitz, J., Hoff, A.L., DeLisi, L.E., and Csernansky, J.G. Correlates of cognitive deficits in first episode schizophrenia. *Schizophr Res* 2004; 68:1–9.
25. Kirkpatrick, B., Buchanan, R.W., Ross, D.E., and Carpenter, W.T. Jr A separate disease within the syndrome of schizophrenia. *Arch Gen Psychiatry* 2001; 58:165–171.
26. Galderisi, S., Maj, M., Mucci, A., Cassano, G.B., Invernizzi, G., Rossi, A., Vita, A., Dell'Osso, L., Daneluzzo, E., and Pini, S. Historical, psychopathological, neurological, and neuropsychological aspects of deficit schizophrenia: a multicenter study. *Am J Psychiatry* 2002; 159:983–990.
27. Buchanan, R.W., Strauss, M.E., Kirkpatrick, B., Holstein, C., Breier, A., and Carpenter, W.T. Neuropsychological impairments in deficit vs. nondeficit forms of schizophrenia. *Arch Gen Psychiatry* 1994; 51:804–811.
28. Cannon, T.D., Zorrilla, L.E., Shtasel, D., Gur, R.E., Gur, R.C., Marco, E.J., Moberg, P., and Price, R.A. Neuropsychological functioning in siblings discordant for schizophrenia and healthy volunteers. *Arch Gen Psychiatry* 1994; 51:651–661.
29. Censits, D.M., Ragland, J.D., Gur, R.C., and Gur, R.E. Neuropsychological evidence supporting a neurodevelopmental model of schizophrenia: a longitudinal study. *Schizophr Res* 1997; 24:289–298.
30. Weickert, T.W., Goldberg, T.E., Marenco, S., Bigelow, L.B., Egan, M.F., and Weinberger, D.R. Comparison of cognitive performances during a placebo period and an atypical antipsychotic treatment period in schizophrenia: critical examination of confounds. *Neuropsychopharmacology* 2003; 28:1491–1500.
31. Gur, R.E., Kohler, C., Ragland, J.D., Siegel, S.J., Bilker, W.B., Loughhead, J., Phend, N., and Gur, R.C. Neurocognitive performance and clinical changes in olanzapine-treated patients with schizophrenia. *Neuropsychopharmacology* 2003; 28:2029–2036.
32. Hagger, C., Buckley, P., Kenny, J.T., Friedman, L., Ubogy, D., and Meltzer, H.Y. Improvement in cognitive functions and psychiatric symptoms in treatment-refractory schizophrenic patients receiving clozapine. *Biol Psychiatry* 1993; 34:702–712.
33. Buchanan, R.W., Holstein, C., and Breier, A. The comparative efficacy and long-term effect of clozapine treatment on neuropsychological test performance. *Biol Psychiatry* 1994; 36:717–725.
34. Grace, J., Bellus, S.B., Raulin, M.L., Herz, M.I., Priest, B.L., Brenner, V., Donnelly, K.L., Smith, P., and Gunn, S. Long-term impact of clozapine and psychosocial treatment on psychiatric symptoms and cognitive functioning. *Psychiatr Serv* 1996; 47:41–45.
35. Hoff, A.L., Faustman, W.O., Wieneke, M., Espinoza, S., Costa, M., Wolkowitz, O., and Csernansky, G.G. The effects of clozapine on symptom reduction, neurocognitive function, and clinical management in treatment-refractory state hospital schizophrenic inpatients. *Neuropsychopharmacology* 1996; 15:361–369.
36. Galletly, C.A., Clark, C.R., McFarlane, A.C., and Weber, D.L. Relationships between changes in symptom ratings, neuropsychological test performance and quality of life in schizophrenic patients treated with clozapine. *Psychiatry Res* 1997; 72:161–166.
37. Bilder, R.M., Goldman, R.S., Volavka, J., Czobor, P., Hoptman, M., Sheitman, B., Lindenmayer, J.P., Citrome, L., McEvoy, J., Kunz, M., Chakos, M., Cooper, T.B., Horowitz, T.L., and Lieberman, J.A. Neurocognitive effects of clozapine, olanzapine, risperidone, and haloperidol in patients with chronic schizophrenia or schizoaffective disorder. *Am J Psychiatry* 2002; 159:1018–1028.
38. Keefe, R.S., Seidman, L.J., Christensen, B.K., Hamer, R.M., Sharma, T., Sitskoorn, M.M., Lewine, R.R., Yurgelun-Todd, D.A., Gur, R.C., Tohen, M., Tollefson, G.D., Sanger, T.M., and Lieberman, J.A. Comparative effect of atypical and conventional antipsychotic drugs on neurocognition in first-episode psychosis: a randomized, double-blind trial of olanzapine versus low doses of haloperidol. *Am J Psychiatry* 2004; 161:985–995.
39. Heaton, R.K., Gladsjo, J.A., Palmer, B.W., Kuck, J., Marcotte, T.D., and Jeste, D.V. Stability and course of neuropsychological deficits in schizophrenia. *Arch Gen Psychiatry* 2001; 58:24–32.
40. Fabrazzo, M., La Pia, S., Monteleone, P., Mennella, R., Esposito, G., Pinto, A., and Maj, M. Fluvoxamine increases plasma and urinary levels of clozapine and its major metabolites in a time- and dose-dependent manner. *J Clin Psychopharmacology* 2000; 20:708–710.
41. Harrigan, E.P., Miceli, J.J., Anziano, R., Watsky, E., Reeves, K.R., Cutler, N.R., Sramek, J., Shiovitz, T., and Middle, M. A randomized evaluation of the effects of six antipsychotic agents on QTc, in the absence and presence of metabolic inhibition. *J Clin Psychopharmacol* 2004; 24(1):62–9.
42. de Leon, J. Psychopharmacology: atypical antipsychotic dosing: the effect of co-medication with anticonvulsants. *Psychiatric Services* 2004; 55:125–128.
43. Goff, D., and Baldessarini, R. Drug interactions with antipsychotic agents. *J Clin Psychopharmacol* 1993; 13:57–67.
44. Roth, B.L., Meltzer, H.Y., and Khan, N. Binding of typical and atypical antipsychotic drugs to multiple neurotransmitter receptors. *Adv Pharmacol* 1998; 42:482–485.
45. Lawler, C.P., Prioleau, C., Lewis, M.M., Mak, C., Jiang, D., Schetz, J.A., Gonzalez, A.M., Sibley, D.R., and Mailman, R.B. Interactions of the novel antipsychotic aripiprazole (OPC-14597) with dopamine and serotonin receptor subtypes. *Neuropsychopharmacology* 1999; 20:612–27.
46. Arnt, J., and Skarsfeldt, T. Do novel antipsychotics have similar pharmacological characteristics? A review of evidence. *Neuropsychopharmacology* 1998; 18:63–101.
47. Bymaster, F., Calligaro, D., Falcone, J., Marsh, R., Moore, N., Tye, N., Seaman, P., and Wong, D. Radioreceptor binding profile of the atypical antipsychotic olanzapine. *Neuropsychopharmacology* 1996; 14:87–96.
48. Heresco-Levy, U., Javitt, D., Ermilov, M., Mordel, C., Silipo, G., and Lichenstein, M. Efficacy of high-dose glycine in the treatment of enduring negative symptoms of schizophrenia. *Arch Gen Psychiatry* 1999; 56:29–36.
49. Goff, D., Tsai, G., Levitt, J., Amico, E., Manoach, D., Schoenfeld, D., Hayden, D., McCarley, R., and Coyle, J. A placebo-controlled trial of D-cycloserine added to conventional neuroleptics in patients with schizophrenia. *Arch Gen Psych* 1999; 56:21–27.

50. Goff, D., Henderson, D., Evins, A., and Amico, E. A placebo-controlled crossover trial of D-cycloserine added to clozapine in patients with schizophrenia. *Biol Psychiatry* 1999; 45: 512–514.
51. Evins, A., Fitzgerald, S., Wine, L., Roselli, R., and Goff, D. A placebo controlled trial of glycine added to clozapine in schizophrenia. *Am J Psychiatry* 2000; 157:826–828.
52. Bymaster, F.P., Felder, C.C., Tzavara, E., Nomikos, G.G., Calligaro, D.O., and McKinzie, D.L. Muscarinic mechanisms of antipsychotic atypicality. *Prog Neuropsychopharmacol Biol Psychiatry* 2003; 27(7):1125–43.
53. Freudenreich, O., and Goff, D.C. Antipsychotic combination therapy in schizophrenia. A review of efficacy and risks of current combinations. *Acta Psychiatr Scand* 2002; 106:323–30.
54. Goff, D., and Baldessarini, R. Antipsychotics, in *Drug Interactions in Psychiatry*. In: Ciraulo, D., Shader, R., Greenblatt, D., and Creelman, W., eds. Baltimore, Maryland: Williams & Wilkins 1995; pp. 129–174.
55. Harvey, P.D., Green, M.F., McGurk, S.R., and Meltzer, H.Y. Changes in cognitive functioning with risperidone and olanzapine treatment: a large-scale, double-blind, randomized study. *Psychopharmacology* 2003; 169:404–411.
56. Green, M.F., Marshall, B.D. Jr., Wirshing, W.C., Ames, D., Marder, S.R., McGurk, S., Kern, R.S., and Mintz, J. Does risperidone improve verbal working memory in treatment-resistant schizophrenia? *Am J Psychiatry* 1997; 154:799–804.
57. Palmer, B.W., Heaton, R.K., Paulsen, J.S., Kuck, J., Braff, D., Harris, M.J., Zisook, S., and Jeste, D.V. Is it possible to be schizophrenic yet neuropsychologically normal? *Neuropsychology* 1997; 11:437–446.
58. Kremen, W.S., Seidman, L.J., Faraone, S.V., Toomey, R., and Tsuang, M.T. 2000; The paradox of normal neuropsychological function in schizophrenia. *Journal of Abnormal Psychology* 109:743–752.
59. Goldberg, T.E., Ragland, J.D., Torrey, E.F., Gold, J.M., Bigelow, L.B., and Weinberger, D.R. Neuropsychological assessment of monozygotic twins discordant for schizophrenia. *Archives of General Psychiatry* 47(11):1066–72 1990;.
60. Green, M.F., Nuechterlein, K.H., Gold, J.M., Barch, D.M., Cohen, J., and Essock, S. et al. Approaching a consensus cognitive battery for clinical trials in schizophrenia: The NIMH-MATRICES conference to select cognitive domains and test criteria. *Biological Psychiatry* 2004; 56:301–7.
61. Laughren, T. A regulatory perspective on psychiatric syndromes in Alzheimer disease. *Am J Geriatr Psychiatry* 2001; Fall 9(4):340–5.
62. Green, M.F., and Nuechterlein, K.H. Should schizophrenia be treated as a neurocognitive disorder? *Schizophrenia Bulletin* 1999; 25:309–319.
63. Smith, T.E., Hull, J.W., Romanelli, S., Fertuck, E., and Weiss, K.A. Symptoms and neurocognition as rate limiters in skills training for psychotic patients. *American Journal of Psychiatry* 1999; 156:1817–1818.
64. Mueser, K.T., Bellack, A.S., Douglas, M.S., and Wade, J.H. Prediction of social skill acquisition in schizophrenic and major affective disorder patients from memory and symptomatology. *Psychiatry Research* 1991; 37:281–296.
65. Leon, A.C. Multiplicity-Adjusted Sample Size Requirements: A Strategy to Maintain Statistical Power when Using the Bonferroni Adjustment. *J Clinical Psychiatry*. (in press).
66. Patterson, T.L., Goldman, S., McKibbin, C.L., Hughs, T., and Jeste, D.V. UCSD performance-based skills assessment: development of a new measure of everyday functioning for severely mentally ill adults. *Schizophrenia Bulletin* 2001; 27(2):235–245.
67. Bellack, A.S., Sayers, M., Mueser, K., and Bennett, M. Evaluation of social problem solving in schizophrenia. *Journal of Abnormal Psychology* 1994; 103:371–378.
68. Klapow, J.C., Evans, J., Patterson, T.L., Heaton, R.K., Koch, W.L., and Jeste, D.V. Direct assessment of functional status in older patients with schizophrenia. *American Journal of Psychiatry* 1997; 154:1022–1024.
69. Addington, J., and Addington, D. Neurocognitive and social functioning in schizophrenia. *Schizophrenia Bulletin* 1999; 25:173–182.
70. Heaton, R.K., Marcotte, T.D., Rivera-Mindt, M., Moore, D.J., Bentley, H., and McCutchan, J.A. et al. The impact of HIV-associated neuropsychological impairment on everyday functioning. *Journal of the International Neuropsychological Society*. in press.
71. Heaton, R.K., Marcotte, T.D., White, D.A., Ross, D., Meredith, K., and Taylor, M.J. et al. The nature and vocational significance of neuropsychological impairment associated with HIV infection. *The Clinical Neuropsychologist* 1996; 10:1–14.
72. Mortiz, S., Ferahli, S., and Naber, D. Memory and attention performance in psychiatric patients: lack of correspondence between clinician-rated and patient-rated functioning with neuropsychological test results. *Journal of the International Neuropsychological Society* in press.
73. van den Bosch, R.J., and Rombouts, R.P. Causal mechanisms of subjective cognitive dysfunction in schizophrenic and depressed patients. *Journal of Nervous and Mental Disease* 1998; 186(6):364–8.
74. van Gorp, W.G., Satz, P., Hinkin, C., Selnes, O., Miller, E.N., and McArthur, J. et al. Metacognition in HIV-1 seropositive asymptomatic individuals: self-ratings versus objective neuropsychological performance. *Multicenter AIDS Cohort Study (MACS)*. *Journal of Clinical and Experimental Neuropsychology* 1991; 13(5):812–9.
75. Horner, M.D., Harvey, R.T., and Denier, C.A. Self-report and objective measures of cognitive deficit in patients entering substance abuse treatment. *Psychiatry Research* 1999; 86:155–61.
76. Hinkin, C.H., van Gorp, W.G., Satz, P., Marcotte, T., Durvasula, R.S., and Wood, S. et al. Actual versus self-reported cognitive dysfunction in HIV-1 infection: memory-metamemory dissociations. *Journal of Clinical and Experimental Neuropsychology* 1996; 18:431–43.
77. Leon, A.C., Marzuk, P.M., and Portera, L. More reliable outcome measures can reduce sample size requirements. *Arch Gen Psychiatry* 1995; 52:867–871.
78. Laughren, T. Comorbid mood disorders and medical illness: a Food and Drug Administration perspective. *Biol Psychiatry* 2003; 54:195–9.
79. International Conference on Harmonisation (ICH). *ICH/FDA: E9 Statistical Principles for Clinical Trials*. www.fda.gov/cder/guidance/ICH\_E9-fnl.PDF
80. Wright, S. The Method of Path Coefficients. *Annals of Mathematical Statistics* 1934; 5:161–215.
81. Cochran, W.G. The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics* 1968; 24:295–313.
82. Fleiss, J.L. *Statistical Methods and Rates and Proportions*. 1981; New York: John Wiley and Sons.

83. Green, M.F., Marder, S.R., Glynn, S.M., McGurk, S.R., Wirshing, W.C., Wirshing, D.A., Liberman, R.P., and Mintz, J. The neurocognitive effects of low-dose haloperidol: a two-year comparison with risperidone. *Biol Psychiatry* 2002; 51:972–8.
84. Tune, L.E., Strauss, M.E., Lew, M.F., Breitlinger, E., and Coyle, J.T. Serum levels of anticholinergic drugs and impaired recent memory in chronic schizophrenic patients. *Am J Psychiatry* 1982; 139:1460–1462.
85. Laird, N.M. Missing data in longitudinal studies. *Statistics in Medicine* 1988; 7:305–315.
86. Goldstein, H. Multilevel mixed linear model analysis using iterative generalised test squares. *Biometrika* 1986; 73:43–56.
87. Bryk, A.S., and Raudenbush, S.W. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 1992; Newbury Park, Calif: Sage.
88. Gibbons, R.D. et al. Some conceptual and statistical issues in analysis of longitudinal psychiatric data. Application to the NIMH treatment of Depression Collaborative Research Program dataset. *Arch Gen Psychiatry* 1993; 50(9):739–50.
89. Hedeker, D., and Gibbons, R.D. A random-effects ordinal regression model for multilevel analysis. *Biometrics* 1994; 50(4):933–944.
90. Diggle, P.J. et al. *Analysis of Longitudinal Data*. 2nd Edition 2002; Oxford: Oxford University Press.
91. Laird, N.M., and Ware, J.H. Random-effects models for longitudinal data. *Biometrics* 1982; 38(4):963–974.
92. Hedeker, D., and Gibbons, R.D. Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods* 1997; 2:64–78.
93. Lavori, P.W. Clinical trials in psychiatry: should protocol deviation censor patient data? *Neuropsychopharmacology* 1992; 6:39–48.