

Functional Neuroimaging and Electrophysiology Biomarkers for Clinical Trials for Cognition in Schizophrenia

Raymond Y. Cho^{1,2}, Judith M. Ford^{3–4}, John H. Krystal^{3–4}, Marc Laruelle⁵, Bruce Cuthbert⁶, and Cameron S. Carter^{7–8}

²Department of Psychiatry, University of Pittsburgh School of Medicine, Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Pittsburgh, PA 15213; ³Department of Psychiatry, Yale University; ⁴Schizophrenia Biological Research Center, Connecticut VA Healthcare System; ⁵Department of Psychiatry, Columbia University; ⁶Adult Psychopathology and Psychosocial Intervention Research, National Institute of Mental Health; ⁷Department of Psychiatry, University of Pittsburgh School of Medicine; ⁸Departments of Psychiatry and Psychology, University of California at Davis

“Measurement and Treatment Research to Improve Cognition in Schizophrenia” or MATRICS is a National Institute of Mental Health (NIMH)–sponsored research effort to drive novel drug development for the treatment of cognitive impairments in schizophrenia. On September 9–10, 2004, MATRICS convened representatives from NIMH, academia, and industry for the final of a series of consensus-oriented conferences, with this meeting entitled “New Approaches to Assessing and Improving Cognition in Schizophrenia.” In this article, we present a summary of the discussions that focused on functional neuroimaging and electrophysiology biomarkers, outlining the research priorities, which included developing further consensus regarding the cognitive processes that are critical to examine in schizophrenia and the paradigms to examine them, resolving issues of test-retest reliability and discrepant findings in the literature, and additional basic research regarding the methods before fully implementing these paradigms in drug trials. Also summarized are the relative advantages and disadvantages of the different approaches and varying perspectives among the participants regarding these tools. Our hope is that this report will help guide and further stimulate coordinated efforts among researchers, industry, and funding agencies in meeting the urgent need for novel pharmacotherapeutics for cognitive impairments in schizophrenia and useful neurobiological

indexes for tracking the resulting improvements in cognitive function.

Key words: EEG/fMRI/psychometrics/
test-retest reliability/multicenter studies

Introduction

“Measurement and Treatment Research to Improve Cognition in Schizophrenia” or MATRICS is a National Institute of Mental Health (NIMH)–sponsored research effort driven by the premise that, despite the significant advances in the studies of cognitive and neural disturbances in schizophrenia over the past few decades, there has been a conspicuous absence in the development and availability of mechanistically novel drugs. On September 9–10, 2004, MATRICS convened representatives from NIMH, academia, and industry in Potomac, Maryland, for the final of a series of consensus-oriented conferences. During this 2-day meeting, there were intensive discussions thematically guided by the overall theme—“New Approaches to Assessing and Improving Cognition in Schizophrenia”—with several breakout groups for more focused discussions. In this article, we present a summary of the discussions that focused on functional neuroimaging and electrophysiology biomarkers, outlining some of the research priorities, relative advantages and disadvantages, and different perspectives among the participants regarding these tools and their roles in the development of novel agents for cognition in schizophrenia.

General Discussion

The discussions began with some consideration of some definitional issues, including what distinguished surrogate markers and biomarkers. While the term *surrogate marker* seemed more clearly to imply any marker that would be acceptable in the place of a primary index of a disturbance, how biomarkers could be defined was more variable. There was agreement that a good biomarker would reveal something about the disease mechanisms and that a desirable property would be that it

¹To whom correspondence should be addressed; tel: (412) 586-9250, fax: (412) 246-5880, e-mail: chory@upmc.edu.

permit more efficient prediction of an end point than the measurement of the end point itself.

However, while it seems that none of the present indexes of cognition-related brain function would, in any strict sense, qualify as a biomarker, it was instructive to realize that there is not necessarily a strong call from the representatives from industry for such markers. Rather, what was articulated—perhaps due, in part, to realistic expectations about what can currently be achieved—was a need for any reliable measures of brain activity associated with cognition that would be sensitive to changes with medications that were predictive of clinical response.

In the following summary of the discussion regarding biomarkers, our remarks will be limited primarily to functional magnetic resonance imaging (fMRI) and electroencephalograph (EEG) approaches, since these were the 2 methods that most participants were most familiar with and reflect what is most commonly used currently in the field. This, of course, is not to downplay the potential importance of other methods such as magnetic resonance spectroscopy, positron emission tomography (PET), or magnetoencephalography (MEG). In fact, the single neuroimaging bioassay that has received broad endorsement by the pharmaceutical industry is the application of PET or single photon emission computed tomography neuroreceptor imaging for the purpose of determining the dose-related occupancy of brain receptor targets for the purpose of guiding clinical dosing. These approaches can also be used to assess the biological effectiveness of drugs by tracking changes in endogenous markers with treatment. However, this review will focus on functional imaging modalities. It will focus on fMRI and EEG approaches because the general guiding principals and applications associated with fMRI apply to functional neuroimaging with PET, with similar parallels between EEG and MEG.

Neuroimaging Biomarkers

There were 3 broad issues that were identified as important. The first concerns identifying the cognitive domains that will be critical to examine and how to examine them. The second concerns establishing test-retest reliability, obviously important to establish for any given experimental paradigm but seldom performed in the field. And finally, there is the important issue of how drugs could affect the neurovasculature and thus the blood oxygen level-dependent (BOLD) signal in fMRI.

Cognitive Domains and Experimental Approaches

Critical to the potential utility of fMRI and EEG for studying drug effects on cognition is a careful consideration of exactly which cognitive domains to study and the experimental designs used to study them—we are oth-

erwise faced with “garbage-in-garbage-out.” If the cognitive constructs and tools employed to probe neural function are not valid or reliable, we will be faced with inconsistent findings and problems with interpretability. An important, though challenging, aspect of this process will be the psychometric norming of the experimental paradigms. Though it is sometimes assumed that cognitive experimental paradigms can be implemented in the scanner without modification, in practice, it can be expected to be an iterative process before the task elicits the desired psychometric and neurometric properties in the MRI environment. This process may also have to be repeated for schizophrenia patients, as the psychometric and neurometric properties of the paradigms may be significantly different for them. And then, as with any norming process, very large, perhaps multicenter, studies will need to be conducted to properly norm the paradigms, with all the inherent scientific, logistical, and funding challenges. One of the practical challenges will be to provide adequate incentives for actually carrying forward the norming process to completion, which, after the initial refinement of the paradigms, has the risk of becoming less scientifically interesting and more tedious and mechanical.

Test-Retest Reliability of fMRI Paradigms

There was consensus that there need to be better methods for establishing the reliability of the fMRI method itself and to resolve some discrepant findings in the literature. Test-retest reliability, though tacitly recognized as important, is almost never performed. There are only a few published studies¹ to serve as demonstrations of the validity of the method, but these studies tend to have a small number of subjects and examine reliability at the region level rather than at the voxel level. These studies also examine limited aspects of the signal, and since fMRI can offer different types of information concerning brain activity, it would be desirable to ascertain what aspects of the signal would be most useful and interesting in assessing drug effects. For instance, will it be percent signal change? The number of activated voxels? Will only thresholded statistical maps be examined? Also, regarding regional differences across the brain, though there was agreement that posterior regions tend to produce more reliable activations, it was also recognized that there would be great value in conducting a systematic investigation of reliability as a function of brain region that, as yet, does not exist in the literature. The Function Biomedical Informatics Research Network (FBIRN, a National Institutes of Health–sponsored multisite effort to coordinate functional neuroimaging investigations of schizophrenia) is currently collecting data that will allow this question to be addressed for activations in parietal, temporal, and frontal lobe regions in both healthy controls and patients with schizophrenia.

Related to the issue of reliability was a discussion of inconsistencies in the literature. For instance, it was noted that tasks like the N-back can show increases or decreases depending on the specific experimental conditions. These “inconsistencies” are showing predictability² and may be informative as to how medications may affect cortical function and cognition. Practitioners of fMRI generally had a good degree of confidence in the method. Where there are discrepant findings, these are assumed or demonstrated to be due to variations in such factors as task design, subject characteristics, and ways of dealing with performance matching issues. However, it was clear that from the standpoint of industry and scientists outside the field, inconsistency in the literature was cause for some uneasiness regarding the methods. The communication of these concerns suggested the need for moving from understanding and explanations of these discrepancies to explicit and definitive reconciling of the inconsistent findings. However, despite the obvious inherent value, it was also recognized that such efforts would pose significant challenges to both scientists and funding agencies, as they would require 1 or more large studies to be able to adequately probe even a subspace of the relevant permutations of different task parameters, subject characteristics, and data-analytic approaches.

Initiatives such as the FBIRN were suggested as a useful point of departure for addressing the nontrivial scientific and logistical challenges inherent in executing such large-scale fMRI projects. The FBIRN is collecting data from a large number of patients and controls on the Sternberg Item Recognition Paradigm, which, like the N-back, activates the dorsolateral prefrontal cortex (DLPFC). With those data, more definitive statements can be made concerning the relationship between memory load and DLPFC activations and how they depend on performance and clinical variables. Though resolving discrepant findings in the literature has not been the focus of the FBIRN, it can serve as a useful model for implementing the large coordinated efforts that may be required to reconcile discrepant findings.

Drug Effects on the Neurovasculature

The third component of the research agenda involves assessing the effects of pharmacologic agents on the coupling of neural activity and hemodynamic response. This would involve some combination of animal models and nonhemodynamic methods such as EEG or MEG and would likely require a case-by-case examination of each candidate drug. Past approaches for measuring the hemodynamic response function (HRF) have relied on simple activation paradigms such as finger tapping, which elicit robust BOLD signals independent of the involvement of higher cognitive functions. More recent approaches enable estimation of the detailed temporal characteristics of the hemodynamic response such as

the delay of HRF,³⁻⁴ which could be applied in controlled designs to assess the effects of various drugs. To distinguish between neural effects and neurovascular coupling effects, fMRI latency data would need to be combined using measures with high temporal resolution such as event-related potential (ERP) and MEG.

Similar approaches could be applied to the study of drugs that have direct effects on the cerebral vasculature. Even subtle effects on global blood flow may be important, as global effects can significantly impact on the ability of cognitive tasks to induce local changes. Methods such as PET or using contrast agents with MRI will be helpful in examining this issue. Animal models may be of particular value in ascertaining the effect of drugs on microvasculature, since PET and MRI are more suited to assessing more gross changes in hemodynamics. Another issue concerns the relative nature of the BOLD signal—there is potentially great value in further developing ways of establishing absolute quantification of perfusion, since drugs may affect not only relative activation but also the baseline activation. Once a systematic approach to examining such neurovascular effects has been validated, each new candidate compound could then be examined for such potential effects before it is brought to the full implementation in cognitive-activation paradigms.

Electrophysiology Biomarkers

Pros and Cons of EEG Approaches

Though it is clear that EEG and fMRI offer 2 complementary approaches, it was useful to review all the merits and disadvantages of EEG. The first, and oft-stated, advantage is that of temporal resolution, allowing the earliest stages of sensory and perceptual information processing, as well as later cognitive and premotor processes to be distinguished precisely in time. Further, overall processing speed can be assessed without contamination by peripheral motor dysfunction. Some additional advantages are that they can be acquired in a more naturalistic setting; they are not subject to neurovascular effects, the importance of which should not be understated as this issue was highlighted as 1 of the main research priorities for pharmacofMRI; and they are relatively low cost. Other suggested advantages of EEG are that it can be acquired relatively quickly and that some components are automatic and accordingly relatively immune to practice effects. Disadvantages, in addition to limited spatial resolution, include the fact that EEG methods have not been developed across the full array of candidate cognitive and task domains under consideration and are relatively limited in measuring sustained processing. In addressing the trade-off between temporal and spatial resolution for EEG and fMRI, some labs with the technical capability have been performing simultaneous EEG and fMRI recording. Another less costly

and technically more accessible approach has been to iterate back and forth between fMRI and EEG studies.

Priorities for Research

Three research priorities were identified for electrophysiology research, with some obvious overlap with the fMRI research priorities. The first concerns the identification of cognitive paradigms and associated ERP components that would be most useful to examine. Related to this is the second research priority, which calls for characterizing the effects of various known agonists and antagonists on ERP components relevant to schizophrenia. The third, paralleling the fMRI priorities, calls for establishing test-retest reliability and a body of normative data. And again, the focus on EEG is not meant to understate the importance of other complementary approaches such as MEG or psychophysiological measures such as prepulse inhibition but, rather, simply reflects the prevalent experience of the discussants and perhaps the field at large.

Identification of Cognitive and EEG Paradigms

As with fMRI paradigms, some more basic work will be required to determine which cognitive domains and associated ERP components will be useful for assessing drug effects, and this will be true for both humans and animal models. Candidate components can come from sensory and early cognitive processing. For instance, mismatch negativity (MMN), an early auditory ERP elicited by deviant auditory stimuli, has some attractive features such as good test-retest reliability⁵ and has been shown to correlate with functional outcome.⁶ These earlier components have the possible advantage of being relatively automatic and less prone to “noise” from top-down cognitive influences or practice effects, thereby potentially being more direct probes of the functional integrity of underlying neural structures. However, although it has long been recognized that impairments in basic sensory processes may contribute to disturbances in higher-order cognitive domains,⁷⁻⁸ the neural architectures that are assayed may not always map cleanly onto the networks that subserve higher-order cognitive constructs of interest. It is also not completely clear how valid the assumption is that early components are independent of control functions, for example, MMN disruption by sleep deprivation.⁹ These earlier components can be contrasted with later components such as the P300, the error-related negativity, or slow waves such as the contingent negative variation, which are thought to reflect some of the critical cognitive disturbances in schizophrenia but are subject to top-down influences such as task instructions or strategy differences.

Although the focus of discussion was on traditional EEG approaches that examine the amplitudes and latencies of ERP components of interest, more recent advances in EEG data-acquisition and analytic methods were given

at least brief mention. Increasingly higher density recordings (up to 256 electrodes) and better algorithms that address the “inverse problem” have allowed improved and more highly constrained source localizations of the ERPs. Advances in MEG have also created new opportunities to achieve better anatomical localization with the temporal resolution of EEG. Investigations of other aspects of the EEG signal are also gaining prominence, such as examining the coherence of activation across brain regions as well as the EEG’s spectral information using techniques such as Fourier or wavelet decompositions.

Pharmacologic Effects on the EEG Signal

Despite some precedent in calling for a systematic investigation of drug effects on ERP components,¹⁰ it was recognized that there is a relative paucity of literature concerning the pharmacologic effects on the candidate ERP components of interest, and this holds true for agents that have already been identified as possible cognitive-enhancing drugs for schizophrenia. There was consensus concerning the need for some basic information regarding how all the agonists and antagonists commonly thought to be relevant to schizophrenia affect these measures. In addition to providing basic insights into the underlying neurophysiology that underlies the respective EEG signals, this information can also serve the practical purpose of informing the development of model systems. For instance, ketamine is well known to reduce the P300 amplitude and could thus be used as a model system to screen a large number of candidate agents by assessing which of them may restore the P300 amplitude.¹¹ Finally, it was also suggested that all these lines of research could benefit by being extended to longitudinal studies of animal models with implantable electrodes, which could offer valuable insights into the neurodevelopmental aspects of schizophrenia and the effects of drug interventions at various stages of the illness.

Validation and Norming of the Paradigm

In alignment with the fMRI research priorities, establishing test-retest reliability and the collection of normative data were deemed important. Though some limited test-retest data are available on components of interest such as the P300,¹² further data would be of obvious value for each of the ERP components and other dependent measures mentioned above. One suggested strategy for improving reliability that has proven useful in some contexts is to employ measures taken over 2 successive days to increase the stability of the measure.

Normative data, which would have great practical relevance for clinical trials or initial screening applications, are also scarce. Normative data that span different populations, genders, age groups, and environments would all be useful when planning and interpreting the results of such studies. Of course, this endeavor would be beset

by most of the difficulties mentioned for norming paradigms for fMRI approaches, perhaps with the notable exception of cost, which is substantially less for EEG approaches. These challenges might be met with an “EBIRN” for ERP and EEG approaches following the FBIRN in its ability to provide the multicenter infrastructure for answering scientific questions that require such large-scale coordinated efforts.

Summary

While neuroimaging and electrophysiology approaches have allowed significant advances in our understanding of cognition and brain functioning, as well as how they are impaired in schizophrenia, these focused discussions were helpful in clarifying how research employing these tools could be further developed for the identification of useful biomarkers for cognition in schizophrenia. Of primary importance will be to develop further consensus regarding the cognitive processes that are critical to examine in schizophrenia and the development of experimental task paradigms to examine them. There was also recognition that the fundamental issue of test-retest reliability needs to be systematically investigated and that discrepant findings in the literature need to be reconciled. Finally, some additional fundamental information regarding the methods will also be necessary for full implementation of these paradigms in drug trials. For instance, in EEG, it would be useful to know how the receptor types thought to be important to the neurochemistry of schizophrenia affect the ERP components of interest; for fMRI, even more basic questions can be asked about drug effects on the neurovasculature.

These research priorities represent a consensus regarding areas for further development and highlight the significant challenges that confront neuroimaging and electrophysiology research in schizophrenia. Our hope is that these recommendations will help guide and further stimulate coordinated efforts among researchers, industry, and funding agencies—as exemplified by MATRICS—in meeting the urgent need for novel pharmacotherapeutics for cognitive impairments in schizophrenia.

Acknowledgments

We would like to thank Gregory McCarthy for his valuable contributions to the discussions at the

MATRICES meeting (Sept. 9-10, 2004, Potomac, MD).

References

1. Noll DC, Genovese CR, Nystrom L, et al. Estimating test-retest reliability in functional MR imaging II: application to motor and cognitive activation studies. *Magnet Reson Med* 1997;38:508–517.
2. Glahn DC, Ragland JD, Abramoff A, et al. Beyond hypo-frontality: a quantitative meta-analysis of functional neuroimaging studies of working memory in schizophrenia. *Hum Brain Mapp* 2005;25(1):60–69.
3. Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 2000;11(6):735–759.
4. Henson RNA, Price CJ, Rugg MD, Turner R, Friston KJ. Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations. *Neuroimage* 2002; 15(1):83–97.
5. Kathmann N, Frodl-Bauch T, Hegerl U. Stability of the mismatch negativity under different stimulus and attention conditions. *Clin Neurophysiol* 1999;110(2):317–323.
6. Light GA, Braff DL. Mismatch negativity deficits are associated with poor functioning in schizophrenia patients. *Arch Gen Psychiat* 2005;62(2):127–136.
7. McGhie A, Chapman J. Disorders of attention and perception in early schizophrenia. *Brit J Med Psychol* 1961; 34:103–116.
8. Venables PH. Sensory aspects of psychopathology. *Proc Annu Meet Am Psychopathol Assoc* 1969;58:132–143.
9. Raz A, Deouell LY, Bentin S. Is pre-attentive processing compromised by prolonged wakefulness? effects of total sleep deprivation on the mismatch negativity. *Psychophysiology* 2001;38(5):787–795.
10. Callaway E. A psychology for psychopharmacologists. In: Barchas JD, Bunney WE, Jr., eds. *Perspectives in Psychopharmacology: A Collection of Papers in Honor of Earl Usdin*. Emeryville, CA: Alibris; 1988:407–415.
11. Umbricht D, Schmid L, Koller RV, Vollenweider FX, Hell D, Javitt DC. Ketamine-induced deficits in auditory and visual context-dependent processing in healthy volunteers: implications for models of cognitive deficits in schizophrenia. *Arch Gen Psychiat* 2000;57(12):1139–1147.
12. Mathalon DH, Ford JM, Pfefferbaum A. Trait and state aspects of P300 amplitude reduction in schizophrenia: a retrospective longitudinal study. *Biol Psychiat* 2000;47(5): 434–449.