

Neuropsychological Testing and Structural Magnetic Resonance Imaging as Diagnostic Biomarkers Early in the Course of Schizophrenia and Related Psychoses

Elissaios Karageorgiou · S. Charles Schulz · Randy L. Gollub · Nancy C. Andreasen · Beng-Choon Ho · John Lauriello · Vince D. Calhoun · H. Jeremy Bockholt · Scott R. Sponheim · Apostolos P. Georgopoulos

Published online: 19 January 2011
© Springer Science+Business Media, LLC (outside the USA) 2011

Abstract Making an accurate diagnosis of schizophrenia and related psychoses early in the course of the disease is important for initiating treatment and counseling patients and families. In this study, we developed classification models for early disease diagnosis using structural MRI (sMRI) and neuropsychological (NP) testing. We used sMRI measurements and NP test results from 28 patients with recent-onset schizophrenia and 47 healthy subjects, drawn from the larger sample of the Mind Clinical Imaging Consortium. We developed diagnostic models based on Linear Discriminant Analysis (LDA) following two approaches; namely, (a) stepwise (STP) LDA on the original measurements, and (b)

LDA on variables created through Principal Component Analysis (PCA) and selected using the Humphrey-Ilgen parallel analysis. Error estimation of the modeling algorithms was evaluated by leave-one-out external cross-validation. These analyses were performed on sMRI and NP variables separately and in combination. The following classification accuracy was obtained for different variables and modeling algorithms. *sMRI only*: (a) STP-LDA: 64.3% sensitivity and 76.6% specificity, (b) PCA-LDA: 67.9% sensitivity and 72.3% specificity. *NP only*: (a) STP-LDA: 71.4% sensitivity and 80.9% specificity, (b) PCA-LDA: 78.5% sensitivity and 91.5% specificity. *Combined sMRI-NP*: (a) STP-LDA:

E. Karageorgiou · S. R. Sponheim · A. P. Georgopoulos (✉)
Brain Sciences Center (11B),
Veterans Affairs Medical Center,
One Veterans Drive,
Minneapolis, MN 55417, USA
e-mail: omega@umn.edu

E. Karageorgiou
e-mail: karag001@umn.edu

S. R. Sponheim
e-mail: sponh001@umn.edu

E. Karageorgiou · A. P. Georgopoulos
Department of Neuroscience,
University of Minnesota Medical School,
Minneapolis, MN, USA

E. Karageorgiou · A. P. Georgopoulos
Department of Neurology,
University of Minnesota Medical School,
Minneapolis, MN, USA

E. Karageorgiou · A. P. Georgopoulos
Center for Cognitive Sciences,
University of Minnesota Medical School,
Minneapolis, MN, USA

S. C. Schulz · S. R. Sponheim · A. P. Georgopoulos
Department of Psychiatry,
University of Minnesota Medical School,
Minneapolis, MN, USA

S. C. Schulz
e-mail: scs@umn.edu

S. C. Schulz · V. D. Calhoun · H. J. Bockholt
Mind Research Network,
Albuquerque, NM, USA

V. D. Calhoun
e-mail: vcalhoun@unm.edu

H. J. Bockholt
e-mail: jbockholt@mrn.org

R. L. Gollub
Department of Psychiatry,
Massachusetts General Hospital and Harvard Medical School,
Charlestown, MA, USA
e-mail: rgollub@partners.org

64.3% sensitivity and 83.0% specificity, (b) PCA-LDA: 89.3% sensitivity and 93.6% specificity. (i) Maximal diagnostic accuracy was achieved by combining sMRI and NP variables. (ii) NP variables were more informative than sMRI, indicating that cognitive deficits can be detected earlier than volumetric structural abnormalities. (iii) PCA-LDA yielded more accurate classification than STP-LDA. As these sMRI and NP tests are widely available, they can increase accuracy of early intervention strategies and possibly be used in evaluating treatment response.

Keywords Schizophrenia · Schizophreniform · Schizoaffective · PCA · LDA · Biomarkers · Neuropsychology · MRI · Cross-validation · Diagnosis · MCIC

Introduction

Making an accurate diagnosis after a first psychotic episode is important for initiating treatment and counseling patients and families (Harrigan et al. 2003). Classification systems for psychiatric disorders are distinguished from the classification of other diseases by their major reliance on behavioral signs and symptoms. Nearly 40 years ago Robins and Guze (1970) outlined a means for developing laboratory tests for mental disorders in order to improve the reliability and validity of classification and move diagnostic assessment beyond clinical description. They suggested that, following clinical description, one should strive to identify reliable and precise laboratory measures to assist in diagnosis. Neuropsychological (NP) measures of cognitive dysfunction are reliable and in most cases are precise throughout disease progression. On the other hand, structural MRI (sMRI) of the brain can generate reliable and precise estimates of brain volumes. Several experimental studies have revealed abnormalities in schizophrenia using NP or sMRI measures (Andreasen et al. 1990; Ballmaier et al. 2008; Hawkins 1999; Hulshoff Pol and Kahn 2008; Leeson et al. 2008; Midorikawa et al. 2008;

Segall et al. 2009; Wang et al. 2008; Watson et al. 1981). Thus, NP and sMRI are suitable candidate laboratory tests for the early identification of individuals with schizophrenia and related psychoses.

Our goal in this study was to develop and evaluate diagnostic models for recent-onset schizophrenia spectrum illnesses based on NP and sMRI variables, in combination and separately. The choice to evaluate recent-onset patients was based on our intent to develop diagnostic tools that can be of practical adjunct use in a real-life first visit to a mental health clinic. To this end, we drew participants from four geographic regions of the United States (Minnesota, New Mexico, Iowa, and Massachusetts).

Previous work by our group demonstrated the potential of applying Linear Discriminant Analysis (LDA) (Green 1978) on accurately diagnosing adult (Georgopoulos et al. 2007) and adolescent (Pardo et al. 2006) schizophrenic patients. To evaluate recent-onset patients in the current study we followed two approaches. In the first approach, based on stepwise (STP) LDA (Dixon 1992), we selected subsets of pertinent variables which would lead to faster models, risking however the rejection of useful information. In the second approach, based on Principal Component Analysis (PCA) (Green 1978) and LDA, we aimed at maximally exploiting data information from our sample, leading to potentially more accurate tests. Through these two approaches we sought to develop NP and sMRI models that could be used as biomarkers for the diagnosis of recent-onset psychotic disorders. In addition, we sought to obtain an unbiased error estimate of future performance of the modeling algorithms using external-cross-validation (Ambrose and McLachlan 2002).

Method

Subject Selection

Subjects were selected from the larger MIND Clinical Imaging Consortium (MCIC) sample (Segall et al. 2009). In

N. C. Andreasen · B.-C. Ho
Department of Psychiatry,
University of Iowa Carver College of Medicine,
Iowa, IA, USA

N. C. Andreasen
e-mail: nancy-andreasen@uiowa.edu

B.-C. Ho
e-mail: beng-ho@uiowa.edu

J. Lauriello
Department of Psychiatry,
University of New Mexico,
Albuquerque, NM, USA
e-mail: JLauriello@salud.unm.edu

V. D. Calhoun
Department of Electrical and Computer Engineering,
University of New Mexico,
Albuquerque, NM, USA

S. R. Sponheim
Mental Health Patient Service Line,
Veterans Affairs Medical Center,
One Veterans Drive,
Minneapolis, MN 55417, USA

S. R. Sponheim
Department of Psychology,
University of Minnesota Medical School,
Minneapolis, MN 55455, USA

the Consortium, prospective recent-onset patients and healthy subjects underwent structured CASH or SCID-I diagnostic interview evaluations by trained assessors (Andreasen et al. 1992; First et al. 2001), sMRI scanning and NP testing. The disease was labeled as of “recent-onset” when any one of the following criteria was present: (a) psychotic symptoms of any duration with a maximum lifetime exposure of 10 mg haloperidol equivalents, (b) psychotic symptoms of no more than 5 years duration, combined with no or minimal antipsychotic exposure (no more than a total lifetime exposure of the equivalent of 10 mg haloperidol), (c) psychotic symptoms of no more than 5 years duration, combined with no more than 12 weeks of antipsychotic exposure, and with no exposure during the previous 3 weeks, or (d) psychotic symptoms of no more than 2 years duration, combined with no more than 6 months of antipsychotic exposure.

Patients had one of the following diagnoses: schizophrenia, schizophreniform disorder, or schizoaffective disorder. We used data from 47 healthy controls and 28 patients ($N_{\text{total}}=75$) in this analysis. (From a total of 84 subjects available, 9 were excluded due to incomplete NP testing.) The demographic characteristics of the two groups are presented in Table 1 (Hollingshead and Redlich 1958). After complete description of the study to the subjects, written informed consent was obtained. Procedures were in accordance with the policies and ethical standards of the institutional review boards at the four sites and the Declaration of Helsinki (World Medical Association General Assembly 2004).

Neuropsychological Assessment

Neuropsychological assessment was based on a battery of tests (Annett 1970; Brandt 1991; Delis et al. 2001; Miller 1990; Reitan 1958; Ruff and Parker 1993; Shallice 1982; Wechsler 1997a, b; Wilkinson 1993) yielding a total of 75 scores, the specifics of which are given in Table 2. None of the NP variables was a composite of others (i.e. sums, averages), which allowed maximum non-overlapping information to be incorporated in the models without adding redundant variables. It is also important to note that none of these 75 variables was used in diagnosing patients. The approximate duration to complete these tests was 3 hours.

sMRI Processing

Cross-site sMRI acquisition calibration and reliability were established in a preceding study using human phantoms following guidelines developed by the Biomedical Informatics Research Network (BIRN) testbed for morphometry (Jovicich et al. 2006, 2009). The MRI acquisition protocol for high-resolution coronal T1-weighted images were

matched across the four different sites (University of Minnesota: Siemens 3 T Trio, University of Iowa: Siemens 1.5 T Sonata, Massachusetts General Hospital and University of New Mexico: Siemens 1.5 T Avanto): 3 T Trio: TR/TE/flip angle=2.53 s/3.79 ms/7°, 1.5 T Sonata: TR/TE/flip angle=20 ms/6 ms/20°, 1.5 T Avanto TR/TE/flip angle 12 ms/4.76 ms/20°, FOV=16×16 cm, 128 contiguous slices (imaging matrix=256×256, in plane resolution=0.625×0.625 mm, slice thickness=1.5 mm), NEX=3. Acquisition time for this Gradient Echo sequence was approximately 20 minutes.

Volumetric measurements of brain areas were obtained from sMRI scans in an automated manner using the Freesurfer software package (<http://surfer.nmr.mgh.harvard.edu>, Version 4.0.1) (Fischl et al. 2002). Minor manual intervention following the standard Freesurfer processing manuals was required for the MRI data of one subject. A total of 95 sMRI variables were used in this analysis (Table 3). The approximate total duration for an MRI acquisition was 40 minutes.

Model Development

Two types of models were developed for discriminating recent-onset patients from healthy controls. Both analyses were based on LDA (Green 1978). The first analysis was applied on the raw data. STP-LDA was performed using Wilk’s Λ as a criterion, with an enter-criterion of $p=0.05$ and a remove-criterion of $p=0.10$. This analysis yielded classification functions from subsets of the original variables. The second analysis involved a dimensionality reduction and was applied on standardized data. For that purpose, raw values were transformed to z-scores (across subjects, within each variable) and then subjected to PCA (Green 1978). An LDA was performed using the first k principal components as predictors; k was determined using Humphrey-Ilgen’s permutation procedure (O’Connor 2000). Finally, in both analyses, prior probabilities for the two groups were set to be equal in the LDA (chance level=50%). This allowed easier, and more accurate, data interpretation since more controls were in the study (47 vs. 28). If priors were not set equal and a model were biased towards controls, the model would have higher overall accuracy and specificity on the specific sample simply by chance (up to $47/75=62.7\%$), though sensitivity would be lower.

Variable Sets Used

The two analyses mentioned in the preceding paragraph were applied under three variable settings: (a) using only NP variables, (b) using only sMRI variables, and (c) using both NP and sMRI variables. This approach allowed us to evaluate how accurate the prediction is by each set of

Table 1 Subject demographics

| Group | Age (years) | | Duration of Diagnosis (years) | | Treatment Duration (years) | | Sex | | Ethnicity | | Race | | | Parental SES ^a | | | | | Highest Education (years) | | Handedness ^b | | | |
|----------|-------------|------|-------------------------------|-----------|----------------------------|----------|------|--------|-----------|--------------|-----------------|-------|-------|---------------------------|---|----|----|---|---------------------------|-------|-------------------------|----|---|---|
| | Mean | SD | Mean | range | Mean | range | Male | Female | Hispanic | Non-Hispanic | Native American | Asian | Black | White | 1 | 2 | 3 | 4 | 5 | Mean | SD | R | A | L |
| Patients | 24.32 | 5.05 | 2.2 | 0.03–12.3 | 0.8 | 0.04–4.0 | 22 | 6 | 5 | 23 | 0 | 1 | 4 | 23 | 1 | 9 | 13 | 5 | 0 | 13.67 | 2.23 | 26 | 1 | 1 |
| Controls | 26.38 | 6.6 | NA | NA | NA | NA | 27 | 20 | 6 | 41 | 2 | 1 | 0 | 44 | 1 | 10 | 34 | 2 | 0 | 14.98 | 1.81 | 43 | 2 | 2 |

^a SES: socioeconomic status as measured by the modified Hollingshead scale. 1 = wealth, prestige; 2 = advanced degree; 3 = white collars, small businesses; 4 = semi-skilled workers, second level of education; 5 = semi-skilled workers, elementary education

^b R = right-handed; L = left-handed; A = ambidextrous

variables separately, and whether their combination might improve or worsen classification.

External Cross-Validation

In addition to developing models that discriminate the two groups, we were interested in evaluating the accuracy of our modeling algorithms on new subjects through cross-validation. For that purpose, we used an external leave-one-out loop. Specifically, before any analysis was performed, a single subject was left out and was used only to cross-validate our results, avoiding selection bias and, thus, overfitting (Ambroise and McLachlan 2002; Duda et al. 2001). This operation was iteratively applied until all subjects were left out (Fig. 1). With respect to PCA-LDA, the z-scores and principal components of each external subject were computed based on coefficients derived from the remaining subjects.

Ultimately, both STP-LDA and PCA-LDA yielded classification tables from which the performance of the two classification methods was evaluated by computing the following measures: overall accuracy of classification, sensitivity, specificity, positive likelihood-ratio, negative likelihood-ratio, χ^2 and the odds-ratio (i.e. strength and direction of the association).

STP-LDA Models for All Subjects

The external cross-validation analyses described in the preceding section were aimed to assess the success (or error) rates of the classification algorithms but cannot provide consistent information concerning the relative importance of the NP and/or sMRI predictors, since those varied depending on the particular subject left out. This information can be gained by performing the same analyses using all subjects. For that purpose, we developed specific models for the STP-LDA using all 75 available subjects under the three settings described above, and evaluated the relative contribution of each variable to the respective model. Each variable's contribution was assessed based on its standardized coefficient. We further investigated how the variables selected in the STP-LDA models relate to the disease using ANOVA with group as a fixed factor and an ANCOVA using sex, age, and intracranial volume as covariates. In the case of the PCA-LDA the analysis described above was not performed since all variables contribute to all principal components and by extension to the LDA.

Software Packages Used

PCA-LDA was performed using the MatLab software package and built-in functions, except for Humphrey-Ilgen

Table 2 Neuropsychological variables used

| | NP Test ^{a,b,c,d} | NP name | Function/Domain assessed |
|----|---------------------------------|-----------------------------------|-----------------------------------|
| 1 | Annett Scale of Hand Preference | Continuous (SUM) | Handedness |
| 2 | WRAT-3 Reading | Total Reading | Reading/Potential or Premorbid IQ |
| 3 | Benton Visual Retention | Error Categories: Omissions | Visual memory |
| 4 | Benton Visual Retention | Error Categories: Distortions | Visual memory |
| 5 | Benton Visual Retention | Error Categories: Perseverations | Visual memory |
| 6 | Benton Visual Retention | Error Categories: Rotations | Visual memory |
| 7 | Benton Visual Retention | Error Categories: Misplacements | Visual memory |
| 8 | Benton Visual Retention | Error Categories: Size Errors | Visual memory |
| 9 | Benton Visual Retention | Error Categories: Left Errors | Visual memory |
| 10 | Benton Visual Retention | Error Categories: Right Errors | Visual memory |
| 11 | Benton Visual Retention | Number Correct Score | Visual memory |
| 12 | Category Fluency | Vegetables | Verbal/Linguistic |
| 13 | Letter Fluency | Starts With 'F' | Verbal/Linguistic |
| 14 | Letter Fluency | Starts With 'A' | Verbal/Linguistic |
| 15 | Letter Fluency | Starts With 'S' | Verbal/Linguistic |
| 16 | Category Fluency | Animals | Verbal/Linguistic |
| 17 | Category Fluency | Fruits | Verbal/Linguistic |
| 18 | WAIS-3 | Block Design Total Raw Score | Spatial motor coordination |
| 19 | WAIS-3 | Letter-Number Total Raw Score | Working memory |
| 20 | WAIS-3 | Vocabulary Total Raw Score | Verbal/Linguistic |
| 21 | WAIS-3 | Similarities Total Raw Score | Verbal abstraction |
| 22 | WMS-3 (Immediate) | Story A Recall Score | Verbal declarative memory |
| 23 | WMS-3 (Immediate) | Story A Thematic Score | Verbal declarative memory |
| 24 | WMS-3 (Immediate) | Story B 1st recall score | Verbal declarative memory |
| 25 | WMS-3 (Immediate) | Story B 1st recall Thematic score | Verbal declarative memory |
| 26 | WMS-3 (Immediate) | Story B 2nd recall score | Verbal declarative memory |
| 27 | WMS-3 (Immediate) | Story B 2nd recall Thematic score | Verbal declarative memory |
| 28 | Trail Making | Trials A: Time | Problem solving/Executive |
| 29 | Trail Making | Trials A: Errors | Problem solving/Executive |
| 30 | Trail Making | Trials B: Time | Problem solving/Executive |
| 31 | Trail Making | Trials B: Errors | Problem solving/Executive |
| 32 | CalCap | SRT Base: Reaction Time | Processing speed & attention |
| 33 | CalCap | CRT Base: Reaction Time | Processing speed & attention |
| 34 | CalCap | CRT Base: True Positive | Processing speed & attention |
| 35 | CalCap | CRT Base: False Positive | Processing speed & attention |
| 36 | CalCap | CRT SEQ1: Reaction Time | Processing speed & attention |
| 37 | CalCap | CRT SEQ1: True Positive | Processing speed & attention |
| 38 | CalCap | CRT SEQ1: False Positive | Processing speed & attention |
| 39 | CalCap | CRT SEQ2: Reaction Time | Processing speed & attention |
| 40 | CalCap | CRT SEQ2: True Positive | Processing speed & attention |
| 41 | CalCap | CRT SEQ2: False Positive | Processing speed & attention |
| 42 | WMS-3 (Delay) | Story A recall score | Verbal declarative memory |
| 43 | WMS-3 (Delay) | Story A Thematic score | Verbal declarative memory |
| 44 | WMS-3 (Delay) | Story B recall score | Verbal declarative memory |
| 45 | WMS-3 (Delay) | Story B Thematic score | Verbal declarative memory |
| 46 | WMS-3 (Delay) | Recognition Total Score | Verbal declarative memory |
| 47 | WMS-3 (Delay) | Percent Retention | Verbal declarative memory |
| 48 | Tower of London | Moves Made: 3 Ring | Problem solving/Executive |
| 49 | Tower of London | Moves Made: 4 Ring | Problem solving/Executive |

Table 2 (continued)

| | NP Test ^{a,b,c,d} | NP name | Function/Domain assessed |
|----|-------------------------------------|----------------------------|------------------------------|
| 50 | Tower of London | Moves Made: 5 Ring | Problem solving/Executive |
| 51 | Tower of London | Avg. Pickup Time: 3 Ring | Problem solving/Executive |
| 52 | Tower of London | Avg. Pickup Time: 4 Ring | Problem solving/Executive |
| 53 | Tower of London | Avg. Pickup Time: 5 Ring | Problem solving/Executive |
| 54 | Tower of London | Avg. Total Time: 3 Ring | Problem solving/Executive |
| 55 | Tower of London | Avg. Total Time: 4 Ring | Problem solving/Executive |
| 56 | Tower of London | Avg. Total Time: 5 Ring | Problem solving/Executive |
| 57 | Tower of London | Total Trials Time: 3 Ring | Problem solving/Executive |
| 58 | Tower of London | Total Trials Time: 4 Ring | Problem solving/Executive |
| 59 | Tower of London | Total Trials Time: 5 Ring | Problem solving/Executive |
| 60 | Tower of London | Avg. Trial Time: 3 Ring | Problem solving/Executive |
| 61 | Tower of London | Avg. Trial Time: 4 Ring | Problem solving/Executive |
| 62 | Tower of London | Avg. Trial Time: 5 Ring | Problem solving/Executive |
| 63 | WMS-3 Faces (Immediate) | Recognition total score | Visual memory |
| 64 | WMS-3 Faces (Delay) | Recognition Total Score | Visual memory |
| 65 | Hopkins verbal learning (Immediate) | Total Raw: Trial 1 | Verbal declarative memory |
| 66 | Hopkins verbal learning (Immediate) | Total Raw: Trial 2 | Verbal declarative memory |
| 67 | Hopkins verbal learning (Immediate) | Total Raw: Trial 3 | Verbal declarative memory |
| 68 | Grooved pegboard | Dominant Hand: # Drops | Fine motor dexterity & speed |
| 69 | Grooved pegboard | Non-dominant Hand: Time | Fine motor dexterity & speed |
| 70 | Grooved pegboard | Non-dominant Hand: # Drops | Fine motor dexterity & speed |
| 71 | Grooved pegboard | Dominant Hand: Time | Fine motor dexterity & speed |
| 72 | Hopkins verbal learning (Delay) | Delay Total Raw | Verbal declarative memory |
| 73 | Hopkins verbal learning (Delay) | Recognition: Hits | Verbal declarative memory |
| 74 | Hopkins verbal learning (Delay) | Recognition: F/A Related | Verbal declarative memory |
| 75 | Hopkins verbal learning (Delay) | Recognition: F/A Unrelated | Verbal declarative memory |

^aWRAT: Wide Range Achievement Test—Third Edition

^bWAIS-3: Wechsler Adult Intelligence Scale—Third Edition

^cWMS-3: Wechsler Memory Scale-III

^dCalCap: California Computerized Assessment Package

analysis which was based on free access code by O'Connor (2000). Regarding STP-LDA this was performed using the SPSS version 15 software package. Special care was taken to remove a subject from the external leave-one-out loop prior to any analysis (Fig. 1), since the SPSS built-in stepwise function selects variables prior to removing a subject and only removes a subject prior to calculating variable coefficients. Settings for each analysis are described in the respective methods sections.

Results

The results of the external cross-validation are presented in Table 4. Ordered contributions of original variables for the proposed final STP-LDA models are presented in Table 5 and their individual relations to the disease group are presented in Table 6.

In general, models developed under PCA-LDA achieved better classification than those under STP-LDA, except for the sMRI-only setting where results were comparable. For the PCA-LDA, NP-only was better than sMRI-only models, while the combination of NP and sMRI variables provided the best classification. In contrast, classification accuracy for STP-LDA was comparable across all combinations of data (NP-only, sMRI-only, NP-sMRI).

The NP variables that were selected through STP-LDA reflected visual memory, problem solving, executive function, and spatial motor abilities. In all those NP tests patients performed worse than controls. Regarding the selected sMRI variables, they were mostly from frontal, temporal and subcortical areas. Of the sMRI variables that had a significant group effect, patients had smaller left precentral brain volume, but larger volume for all the other structures. This remained consistent

Table 3 sMRI variables used

| | sMRI name | | sMRI name | | sMRI name |
|----|----------------------------------|----|---------------------------------|----|----------------------------------|
| 1 | Right caudal anterior cingulate | 33 | Right transverse temporal | 65 | Left temporal pole |
| 2 | Right caudal middle frontal | 34 | Left caudal anterior cingulate | 66 | Left transverse temporal |
| 3 | Right corpus callosum | 35 | Left caudal middle frontal | 67 | Left Cerebral White Matter |
| 4 | Right cuneus | 36 | Left corpus callosum | 68 | Left Cerebral Cortex |
| 5 | Right entorhinal | 37 | Left cuneus | 69 | Left Lateral Ventricle |
| 6 | Right fusiform | 38 | Left entorhinal | 70 | Left Inferior Lateral Ventricle |
| 7 | Right inferior parietal | 39 | Left fusiform | 71 | Left Cerebellum White Matter |
| 8 | Right inferior temporal | 40 | Left inferior parietal | 72 | Left Cerebellum Cortex |
| 9 | Right isthmus cingulate | 41 | Left inferior temporal | 73 | Left Thalamus Proper |
| 10 | Right lateral occipital | 42 | Left isthmus cingulate | 74 | Left Caudate |
| 11 | Right lateral orbitofrontal | 43 | Left lateral occipital | 75 | Left Putamen |
| 12 | Right lingual | 44 | Left lateral orbitofrontal | 76 | Left Pallidum |
| 13 | Right medial orbitofrontal | 45 | Left lingual | 77 | Third Ventricle |
| 14 | Right middle temporal | 46 | Left medial orbitofrontal | 78 | Fourth Ventricle |
| 15 | Right parahippocampal | 47 | Left middle temporal | 79 | Left Hippocampus |
| 16 | Right paracentral | 48 | Left parahippocampal | 80 | Left Amygdala |
| 17 | Right pars opercularis | 49 | Left paracentral | 81 | Left Accumbens area |
| 18 | Right pars orbitalis | 50 | Left pars opercularis | 82 | Right Cerebral White Matter |
| 19 | Right pars triangularis | 51 | Left pars orbitalis | 83 | Right Cerebral Cortex |
| 20 | Right pericalcarine | 52 | Left pars triangularis | 84 | Right Lateral Ventricle |
| 21 | Right postcentral | 53 | Left pericalcarine | 85 | Right Inferior Lateral Ventricle |
| 22 | Right posterior cingulate | 54 | Left postcentral | 86 | Right Cerebellum White Matter |
| 23 | Right precentral | 55 | Left posterior cingulate | 87 | Right Cerebellum Cortex |
| 24 | Right precuneus | 56 | Left precentral | 88 | Right Thalamus Proper |
| 25 | Right rostral anterior cingulate | 57 | Left precuneus | 89 | Right Caudate |
| 26 | Right rostral middle frontal | 58 | Left rostral anterior cingulate | 90 | Right Putamen |
| 27 | Right superior frontal | 59 | Left rostral middle frontal | 91 | Right Pallidum |
| 28 | Right superior parietal | 60 | Left superior frontal | 92 | Right Hippocampus |
| 29 | Right superior temporal | 61 | Left superior parietal | 93 | Right Amygdala |
| 30 | Right supramarginal | 62 | Left superior temporal | 94 | Right Accumbens area |
| 31 | Right frontal pole | 63 | Left supramarginal | 95 | Intracranial Volume |
| 32 | Right temporal pole | 64 | Left frontal pole | | |

even when accounting for age, sex, and intracranial volume for variables with significant differences between groups.

Discussion

Overall Diagnostic Usefulness

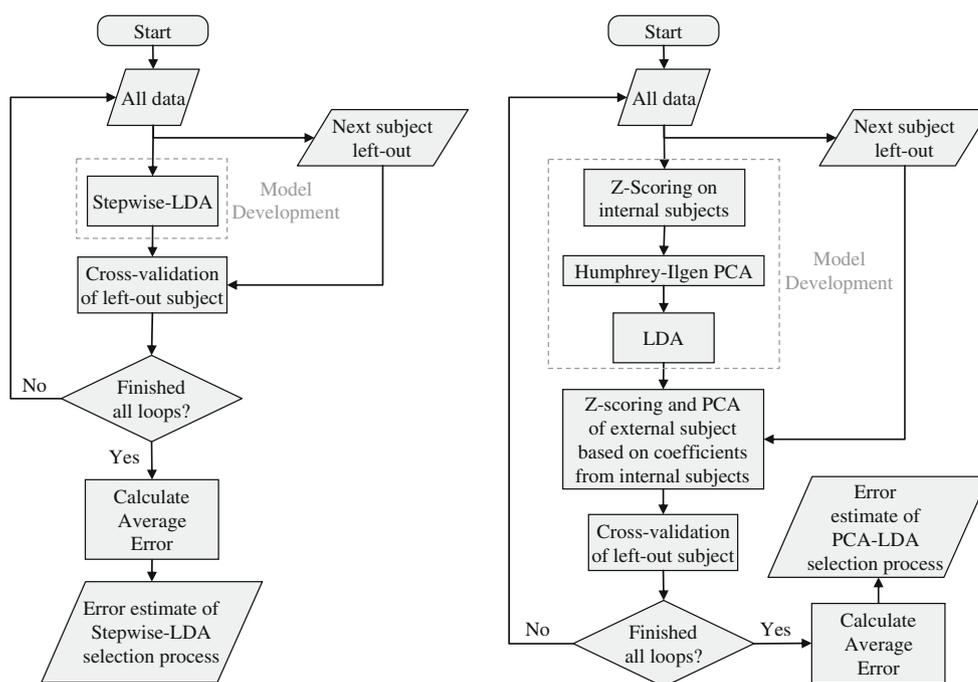
We applied STP-LDA and PCA-LDA under three settings (NP-only, sMRI-only, Combined NP-sMRI) to discriminate recent-onset psychotic patients and controls. Rejection of the null hypothesis that classification was due to chance was achieved for all modeling approaches under all settings ($\chi^2 \geq 11.6$; $p \leq 7 \times 10^{-4}$). These results indicate that both methods

under all three settings can contribute at a variable level to disease diagnosis (i.e. accurately identifying whether an individual does or does not suffer from a psychotic disorder).

PCA-LDA vs. STP-LDA

When comparing between the two modeling methods, there was generally better classification with the PCA-LDA than with the STP-LDA algorithm, with the exception of the sMRI-only setting where results were comparable (compare top vs. bottom of Table 4). This finding indicates that PCA-LDA is more robust as a classification model probably due to the greater amount of accounted total variance. In the case of the STP-LDA, only certain variables are selected each time and, in addition, these variables partially covary

Fig. 1 Flowchart of model development and external cross-validation for the STP-LDA and the PCA-LDA



(i.e. carry common information). On the other hand, the variables selected through PCA are composites of all original variables, and are between them orthogonal (zero covariance).

NP-Only vs. sMRI-only vs. Combined NP-sMRI

Comparison of the three variable settings is slightly more complicated. When comparing individually the NP-only

Table 4 Results of external cross-validation

| | | NP-only | | sMRI-only | | Combined NP-sMRI | |
|---------|--------------------------------|-------------------------------|----|------------------------------|----|------------------------------|----|
| | | SZ | HC | SZ | HC | SZ | HC |
| STP-LDA | Predicted SZ | 20 | 9 | 18 | 11 | 18 | 8 |
| | Predicted HC | 8 | 38 | 10 | 36 | 10 | 39 |
| | Sensitivity (CI)% | 71.4 (51.3–86.8) | | 64.3 (44.1–81.4) | | 64.3 (44.1–81.4) | |
| | Specificity (CI)% | 80.9 (66.7–90.9) | | 76.6 (62.0–87.7) | | 83.0 (69.2–92.4) | |
| | Accuracy (CI)% | 77.3 (66.2–86.2) | | 72.0 (60.4–81.8) | | 76.0 (64.8–85.1) | |
| | Positive likelihood ratio (CI) | 3.73 (2.10–6.41) | | 2.75 (1.57–4.65) | | 3.78 (2.00–7.12) | |
| | Negative likelihood ratio (CI) | 0.35 (0.21–0.59) | | 0.47 (0.29–0.73) | | 0.43 (0.28–0.66) | |
| | Odds ratio (CI) | 10.56 (3.58–31.14) | | 5.89 (2.17–16.25) | | 8.77 (3.01–25.59) | |
| | χ^2 (p) | 20.22 (7×10^{-6}) | | 12.36 (4×10^{-4}) | | 17.31 (3×10^{-5}) | |
| PCA-LDA | Predicted SZ | 22 | 4 | 19 | 13 | 25 | 3 |
| | Predicted HC | 6 | 43 | 9 | 34 | 3 | 44 |
| | Sensitivity (CI)% | 78.6 (59.1–91.7) | | 67.9 (47.7–84.1) | | 89.3 (71.8–97.8) | |
| | Specificity (CI)% | 91.5 (79.6–97.6) | | 72.3 (57.4–84.4) | | 93.6 (82.5–98.7) | |
| | Accuracy (CI)% | 86.7 (76.8–93.4) | | 70.7 (59.0–80.6) | | 92.0 (83.4–97.0) | |
| | Positive likelihood ratio (CI) | 9.23 (4.16–21.55) | | 2.45 (1.47–3.87) | | 13.99 (6.02–32.00) | |
| | Negative likelihood ratio (CI) | 0.23 (0.14–0.40) | | 0.44 (0.26–0.73) | | 0.11 (0.05–0.25) | |
| | Odds ratio (CI) | 39.42 (10.35–149.36) | | 5.52 (2.02–15.10) | | 122.22 (23.99–622.85) | |
| | χ^2 (p) | 38.03 (7×10^{-10}) | | 11.6 (7×10^{-4}) | | 51.5 (7×10^{-13}) | |

CI, 95% confidence intervals for sensitivity, specificity and accuracy calculated based on the binomial theorem
 p, p-values

Table 5 Variable selection and ordered contribution during STP-LDA. All-subject model

| Rank | NP name | sMRI name | NP and sMRI name |
|------|--|--|---|
| 1 | Benton Visual Retention: Error Categories: Left Errors | Frontal: Left precentral | Frontal: Left precentral |
| 2 | Grooved pegboard: Dominant Hand: Time | Subcortical: Right Pallidum | Tower of London: Total Trials Time: 5 Ring |
| 3 | WMS-3 (Delay): Story B recall score | Subcortical: Third Ventricle | WMS-3 (Delay): Recognition Total Score |
| 4 | Tower of London: Total Trials Time: 3 Ring | Frontal: Right frontal pole | Frontal: Left superior frontal |
| 5 | WMS-3 (Delay): Recognition Total Score | Temporal: Right parahippocampal | Subcortical: Left Putamen |
| 6 | Benton Visual Retention: Error Categories: Rotations | | Benton Visual Retention: Error Categories: Distortions |
| 7 | Tower of London: Total Trials Time: 5 Ring | | WMS-3 (Immediate): Story B 1st recall Thematic score |
| 8 | WMS-3 (Delay): Story B Thematic score | | WMS-3 (Delay): Story B recall score |
| 9 | Benton Visual Retention: Error Categories: Perseverations | | Subcortical: Right Accumbens area |
| 10 | Benton Visual Retention: Error Categories: Misplacements | | Frontal: Right superior frontal |
| 11 | Grooved pegboard: Dominant Hand: # Drops | | Benton Visual Retention: Error Categories: Misplacements |
| 12 | WAIS-3: Block Design Total Raw Score | | Grooved pegboard: Non-dominant Hand: Time |
| 13 | | | Temporal: Right temporal pole |

Table 6 Relationship to disease of individual variables from the STP-LDA

| Variable name | Patients vs. Controls | F _[1,73] | p |
|-----------------------------------|-----------------------|---------------------|---------------------|
| Story B recall score | Worse | 79.63 | 3×10^{-13} |
| Story B Thematic score | Worse | 53.22 | 3×10^{-10} |
| Error Categories: Left Errors | Worse | 38.18 | 3×10^{-8} |
| Story B 1st recall Thematic score | Worse | 37.93 | 4×10^{-8} |
| Non-dominant Hand: Time | Worse | 31.54 | 3×10^{-7} |
| Total Trials Time: 5 Ring | Worse | 30.88 | 4×10^{-7} |
| Recognition Total Score | Worse | 24.39 | 5×10^{-6} |
| Error Categories: Distortions | Worse | 24.34 | 5×10^{-6} |
| Dominant Hand: Time | Worse | 22.24 | 10^{-5} |
| Block Design Total Raw Score | Worse | 16.36 | 10^{-4} |
| Error Categories: Misplacements | Worse | 15.99 | 2×10^{-4} |
| Error Categories: Perseverations | Worse | 15.84 | 2×10^{-4} |
| Right frontal pole | Larger | 11.5 | 0.001 |
| Right parahippocampal | Larger | 8.14 | 0.006 |
| Left precentral | Smaller | 7.56 | 0.008 |
| Left Putamen | Larger | 6.66 | 0.012 |
| Total Trials Time: 3 Ring | Worse | 6.03 | 0.016 |
| Right Pallidum | Larger | 4.34 | 0.041 |
| Third Ventricle | Larger | 4.07 | 0.047 |
| Right temporal pole | Larger | 3.45 | 0.067 (ns) |
| Right Accumbens area | Larger | 3.36 | 0.071 (ns) |
| Dominant Hand: # Drops | Worse | 2.03 | 0.158 (ns) |
| Left superior frontal | Larger | 0.25 | 0.615 (ns) |
| Right superior frontal | Larger | 0.21 | 0.651 (ns) |
| Error Categories: Rotations | Worse | 0.03 | 0.858 (ns) |

setting to the sMRI-only setting, better classification is achieved using only NP variables under both algorithms. This finding indicates that there is more pertinent information in the NP relative to sMRI variables and, by extension, that cognitive deficits are probably more sensitive to the occurrence of psychotic disorders than structural abnormalities. However, this finding does not necessitate that an sMRI-only model cannot be developed, but rather that the amount of information used and the approach followed did not lead to any strong sMRI diagnostic measure.

The improved classification of the NP-sMRI combination only for the PCA-LDA, but not for the STP-LDA can be explained as follows. Through PCA-LDA all variables contribute to each principal component, and thus more of the original variance is incorporated in a model. On the other hand, through STP-LDA, a selection always has to be performed and the amount of variance explained will probably not be as large. In our case, the extra variance incorporated in the PCA-LDA was relevant to discriminating healthy subjects from patients. However, if the combination of variables added information irrelevant to disease diagnosis, then classification could have been worse. It is the relevance of variables to disease diagnosis that makes the NP-sMRI PCA-LDA method the best approach.

Best Model

From the modeling methods applied, the best diagnostic outcome for all measures was provided under the Combined NP-sMRI setting of the PCA-LDA algorithm. The high values of sensitivity (89.3%), specificity (93.6%), positive likelihood-ratio (13.99), and diagnostic odds-ratio (122.22), and the low negative likelihood-ratio (0.11) suggest a diagnostic usefulness of the PCA-LDA under combined NP-sMRI. The fact that all variables need to be used in this model makes it a better candidate as an adjunct in disease diagnosis rather than a fast screening tool (i.e. approximately three and a half hours for a subject to undergo all NP and sMRI testing).

Utility of Other Models

On the other hand, when it comes to a faster model, it is more practical to use fewer variables. That can be achieved through the STP-LDA models. In these models, sensitivity and specificity are good, yet not as high as in the best model presented above. Overall, the findings on the STP-LDA models do not make them ideal screening tests, but there are indications of their usefulness as faster tools (duration of about 1 h for the NP-only) and with the introduction of more subjects in the MCIC database prediction could improve. NP-only models can be useful to a health provider since NP tests are straightforward in their administration and, even more,

there are now objective computerized tests that can be widely available to the community, even in remote areas. Similarly, in those cases where patients visiting the doctor's office have already had an MRI, an sMRI-only model could be used as an adjunct during initial evaluation.

Variable Contribution in STP-LDA

The fact that differences in NP score means were more significant than sMRI variables indicates that cognitive deficits are more sensitive for early diagnosis than volumetric structural deficits and, thus, also supports the findings discussed above. The specific variables that were selected by the STP-LDA related to memory (verbal and visual), motor dexterity and speed, visuospatial abilities, and executive function. Patients consistently performed worse in these tests relative to controls (see Table 6). These findings are in agreement with a large body of evidence on the nature of cognitive deficits in chronic and first-episode schizophrenia (Hawkins 1999; Leeson et al. 2008; Midorikawa et al. 2008; Watson et al. 1981).

The sMRI structures that were selected by STP-LDA were frontal, temporal and subcortical. The involvement of basal ganglia in schizophrenia has been reported in MRI studies of minimally medicated patients, however volumetric differences between groups have not been consistently observed, with larger, smaller, or equal volumes in patients versus controls (Ballmaier et al. 2008; Glenthøj et al. 2007; Kawasaki et al. 2004; Velakoulis et al. 2002; Wang et al. 2008). Our findings indicate a tendency for increased putamen and pallidum volumes in patients. On the other hand, the relation observed between increased ventricle size and schizophrenia is better characterized in the literature and consistently observed (Andreasen et al. 1990; Hulshoff Pol and Kahn 2008). Regarding the volumetric differences in the frontal and temporal lobes, results from previous research indicate an association between decreased volume and schizophrenia (Borgwardt et al. 2007; Hazlett et al. 2008; Kawasaki et al. 2004; Velakoulis et al. 2002). However, there are several reports for regional variability within lobes (Selemon 2001). Szeszko et al. (1999) report volume variability within lobes, specifically, more frontal subregions in the frontal lobe having increased size and posterior subregions having decreased size in patients, which agrees with our findings (increased frontal pole volume and decreased precentral volume). The only finding in our results that has not been reported in previous studies is the increased size of the right parahippocampal cortex (Baaré et al. 2001). This could be due to the fact that the patients in the study were recent-onset, and decreased volume of temporal structures is associated with increased disease duration (Velakoulis et al. 2002). On a cellular level, the volumetric differences above can be explained

through synaptic pruning in the cortex (Harrison 1999; Keshavan et al. 1994; Selemon 2001) and an increase in the number of synapses in the basal ganglia (Roberts et al. 2008).

This research, which uses computerized assessment of MRI scans, a thorough neuropsychological test battery, and a sophisticated analysis procedure carries potential for clinical application. It is possible that the MRI and neuropsychological indices may be useful for cases that are difficult to diagnose; however, in the current study only research subjects who met diagnostic criteria on a standard structured interview were included in the analysis. Therefore, the potential for neuropsychological and MRI procedures to assist specifically in complex cases is not fully tested. The present findings represent a first step in identifying indices that are viable candidates for biomarkers to enhance diagnostic determinations in schizophrenia.

As this study addresses the utilization of sMRI and neuropsychological testing for classification, it is appropriate to note prior results in this sample from more traditional group-based analyses. Regarding neuropsychological testing, Sponheim et al. (2009) have reported that both the recent-onset and chronic schizophrenic subjects exhibit a wide range of cognitive deficits compared to controls, which is consistent with the body of findings documenting cognitive dysfunction in schizophrenia. Regarding sMRI, one published report describes gray matter in MCIC subjects combined with a sample of BIRN subjects in order to increase power in detecting differences between patients and controls (Segall et al. 2009). Statistically significant differences in gray matter concentration were evident between schizophrenic patients and controls for similar brain regions across the various study sites. Thus, traditional group-based analyses on sMRI and neuropsychological measures carry some diagnostic utility. Of note, however, the above studies included both recent-onset and chronic schizophrenic patients in their analyses.

Practical Considerations

It is important to keep in mind that we evaluated minimally medicated patients who resemble patients coming to the office for the first time. This makes the derived models useful adjuncts in verifying the diagnosis early in the course of the disease with high specificity and sensitivity. The use of sMRI and NP tests allows objective measures to be incorporated into disease diagnosis, thus negating the subjectivity of structured interviews while allowing, at the same time, non-specialists to contribute in screening for disease through computerized tests. This would be an asset in rural areas and primary care centers. Even more, such models can potentially be used as surrogate biomarkers for treatment response. Full description of the Mind Clinical

Imaging Consortium data base and tools for data abstraction can be accessed as referenced below (see Information Sharing Statement). The databases are currently being expanded to include data on genetic variation in the sample. Analyses that examine a combination of single nucleotide polymorphisms (SNPs), structural MRI, functional MRI, and cognitive performance may further enhance the utility of biomarkers to assist with diagnostic determinations of schizophrenia.

The best model proposed is that of the PCA-LDA with all NP and sMRI variables included. However, useful information can be obtained from the NP variables only, if, for example, a faster test were desirable. In general, the accuracy of modeling algorithms is likely to be improved with increase in the number of subjects studied. Moreover, this approach holds promise for differential diagnosis, for example, between schizophrenia and bipolar disorder (Pardo et al. 2006), with and without psychotic features, as well as between schizophrenia and schizophreniform disorder. Finally, current criteria require 6 months for the specific diagnosis of schizophrenia to be established. If our models successfully pass all phases of biomarker development (Pepe 2005), they could help make these criteria more efficient and lead to early discrimination between schizophreniform disorder and schizophrenia.

Information Sharing Statement

For more information on the methods, models, applications and tools used in this article, when they are publically available, please visit <https://portal.mind.unm.edu/mcic/public/>. Data will be accessible on Xnat (<http://www.xnat.org/>) in the near future.

Acknowledgments Supported by the United States Department of Energy under Award Number DE-FG02-99ER62764 to The Mind Research Network (formerly The Mental Illness and Neuroscience Discovery [MIND] Institute), the Department of Veterans Affairs, the American Legion Brain Sciences Chair, and the Stanley Medical Research Institute.

The authors thank Dr. Christopher Bingham for advice on the PCA, and Mary A. Jacintha and Kathleen E. Kelly for help in study coordination.

References

- Ambrose, C., & McLachlan, G. J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 6562–6566.
- Andreasen, N. C., Ehrhardt, J. C., Swayze, V. W., 2nd, Alliger, R. J., Yuh, W. T., Cohen, G., et al. (1990). Magnetic resonance imaging of the brain in schizophrenia. The pathophysiologic significance of structural abnormalities. *Archives of General Psychiatry*, 47, 35–44.

- Andreasen, N. C., Flaum, M., & Arndt, S. (1992). The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Archives of General Psychiatry*, *49*, 615–623.
- Annett, M. (1970). A classification of hand preference by association analysis. *British Journal of Psychology (London, England: 1953)*, *61*, 303–321.
- Baaré, W. F., van Oel, C. J., Hulshoff Pol, H. E., Schnack, H. G., Durston, S., Sitskoorn, M. M., et al. (2001). Volumes of brain structures in twins discordant for schizophrenia. *Archives of General Psychiatry*, *58*, 33–40.
- Ballmaier, M., Schlagenhaut, F., Toga, A. W., Gallinat, J., Koslowski, M., Zoli, M., et al. (2008). Regional patterns and clinical correlates of basal ganglia morphology in non-medicated schizophrenia. *Schizophrenia Research*, *106*, 140–147.
- Borgwardt, S. J., Riecher-Rossler, A., Dazzan, P., Chitnis, X., Aston, J., Drewe, M., et al. (2007). Regional gray matter volume abnormalities in the at risk mental state. *Biological Psychiatry*, *61*, 1148–1156.
- Brandt, J. (1991). The Hopkins verbal learning test: development of a new memory test with six equivalent forms. *Clinical Neuropsychologist*, *5*, 125–142.
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (DKEFS): Examiner's manual*. San Antonio, TX: The Psychological Corporation.
- Dixon, W. J. (1992). *BMDP statistical software manual*. Berkeley: University of California Press.
- Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (p. 654). New York: Wiley.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. (2001). *Structural clinical interview for DSM-IV-TR axis I disorders, research version, patient edition with psychotic screen (SCID-I/P W/PSY SCREEN)*. New York: New York State Psychiatric Institute, Biometrics Research.
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., et al. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*, 341–355.
- Georgopoulos, A. P., Karageorgiou, E., Leuthold, A. C., Lewis, S. M., Lynch, J. K., Alonso, A. A., et al. (2007). Synchronous neural interactions assessed by magnetoencephalography: a functional biomarker for brain disorders. *Journal of Neural Engineering*, *4*, 349–355.
- Glenthøj, A., Glenthøj, B. Y., Mackeprang, T., Pagsberg, A. K., Hemmingsen, R. P., Jernigan, T. L., et al. (2007). Basal ganglia volumes in drug-naïve first-episode schizophrenia patients before and after short-term treatment with either a typical or an atypical antipsychotic drug. *Psychiatry Research*, *154*, 199–208.
- Green, P. E. (1978). *Analyzing multivariate data*. Hinsdale: The Dryden.
- Harrigan, S. M., McGorry, P. D., & Krstev, H. (2003). Does treatment delay in first-episode psychosis really matter? *Psychological Medicine*, *33*, 97–110.
- Harrison, P. J. (1999). The neuropathology of schizophrenia. A critical review of the data and their interpretation. *Brain: A Journal of Neurology*, *122*(Pt 4), 593–624.
- Hawkins, K. A. (1999). Memory deficits in patients with schizophrenia: preliminary data from the Wechsler Memory Scale-Third Edition support earlier findings. *Journal of Psychiatry & Neuroscience: JPN*, *24*, 341–347.
- Hazlett, E. A., Buchsbaum, M. S., Haznedar, M. M., Newmark, R., Goldstein, K. E., Zelmanova, Y., et al. (2008). Cortical gray and white matter volume in unmedicated schizotypal and schizophrenia patients. *Schizophrenia Research*, *101*, 111–123.
- Hollingshead, A. B., & Redlich, F. C. (1958). *Social class and mental illness; a community study* (p. 442). New York: Wiley.
- Hulshoff Pol, H. E., & Kahn, R. S. (2008). What happens after the first episode? A review of progressive brain changes in chronically ill patients with schizophrenia. *Schizophrenia Bulletin*, *34*, 354–366.
- Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., et al. (2006). Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, *30*, 436–443.
- Jovicich, J., Czanner, S., Han, X., Salat, D., van der Kouwe, A., Quinn, B., et al. (2009). MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *NeuroImage*, *46*, 177–192.
- Kawasaki, Y., Suzuki, M., Nohara, S., Hagino, H., Takahashi, T., Matsui, M., et al. (2004). Structural brain differences in patients with schizophrenia and schizotypal disorder demonstrated by voxel-based morphometry. *European Archives of Psychiatry and Clinical Neuroscience*, *254*, 406–414.
- Keshavan, M. S., Anderson, S., & Pettegrew, J. W. (1994). Is schizophrenia due to excessive synaptic pruning in the prefrontal cortex? The Feinberg hypothesis revisited. *Journal of Psychiatric Research*, *28*, 239–265.
- Leeson, V. C., Barnes, T. R., Harrison, M., Matheson, E., Harrison, I., Mutsatsa, S. H., et al. (2008). The relationship between IQ, memory, executive function, and processing speed in recent-onset psychosis: 1-year stability and clinical outcome. *Schizophrenia bulletin*, [epub ahead of print]
- Midorikawa, A., Hashimoto, R., Noguchi, H., Saitoh, O., Kunugi, H., & Nakamura, K. (2008). Impairment of motor dexterity in schizophrenia assessed by a novel finger movement test. *Psychiatry Research*, *159*, 281–289.
- Miller, E. N. (1990). *California computerized assessment battery (CalCAP) manual*. Los Angeles: Norland Software.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc.*, *32*, 396–402.
- Pardo, P. J., Georgopoulos, A. P., Kenny, J. T., Stuve, T. A., Findling, R. L., & Schulz, S. C. (2006). Classification of adolescent psychotic disorders using linear discriminant analysis. *Schizophrenia Research*, *87*, 297–306.
- Pepe, M. S. (2005). Evaluating technologies for classification and prediction in medicine. *Statistics in Medicine*, *24*, 3687–3696.
- Reitan, R. M. (1958). Validity of the trail making test as indicator of organic brain damage. *Perceptual and Motor Skills*, *8*, 271–276.
- Roberts, R. C., Roche, J. K., & Conley, R. R. (2008). Differential synaptic changes in the striatum of subjects with undifferentiated versus paranoid schizophrenia. *Synapse (New York, N.Y.)*, *62*, 616–627.
- Robins, E., & Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *The American Journal of Psychiatry*, *126*, 983–987.
- Ruff, R. M., & Parker, S. B. (1993). Gender- and age-specific changes in motor speed and eye-hand coordination in adults: normative values for the Finger Tapping and Grooved Pegboard Tests. *Perceptual and Motor Skills*, *76*, 1219–1230.
- Segall, J. M., Turner, J. A., van Erp, T. G., White, T., Bockholt, H. J., Gollub, R. L., et al. (2009). Voxel-based morphometric multisite collaborative study on schizophrenia. *Schizophrenia Bulletin*, *35*, 82–95.
- Selemon, L. D. (2001). Regionally diverse cortical pathology in schizophrenia: clues to the etiology of the disease. *Schizophrenia Bulletin*, *27*, 349–377.
- Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *298*, 199–209.
- Sponheim, S. R., Jung, R. E., Seidman, L. J., Mesholam-Gately, R. I., Manoach, D. S., O'Leary, D. S., et al. (2009). Cognitive deficits

- in recent-onset and chronic schizophrenia. *Journal of Psychiatric Research*, 44, 421–8.
- Szeszko, P. R., Bilder, R. M., Lencz, T., Pollack, S., Alvir, J. M., Ashtari, M., et al. (1999). Investigation of frontal lobe subregions in first-episode schizophrenia. *Psychiatry Research*, 90, 1–15.
- Velakoulis, D., Wood, S. J., Smith, D. J., Soulsby, B., Brewer, W., Leeton, L., et al. (2002). Increased duration of illness is associated with reduced volume in right medial temporal/anterior cingulate grey matter in patients with chronic schizophrenia. *Schizophrenia Research*, 57, 43–49.
- Wang, L., Mamah, D., Harms, M. P., Karnik, M., Price, J. L., Gado, M. H., et al. (2008). Progressive deformation of deep brain nuclei and hippocampal-amygdala formation in schizophrenia. *Biological Psychiatry*, 64, 1060–1068.
- Watson, C. G., Gasser, B., Schaefer, A., Buranen, C., & Wold, J. (1981). Separation of brain-damaged from psychiatric patients with ability and personality measures. *Journal of Clinical Psychology*, 37, 347–353.
- Wechsler, D. (1997a). *Wechsler adult intelligence scale*. San Antonio: Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler memory scale*. San Antonio: Psychological Corporation.
- Wilkinson, G. (1993). *Wide range achievement test*. Wilmington: Wide Range, Inc.
- World Medical Association General Assembly. (2004). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Journal Internationale de Bioéthique*, 15, 124–129.