Title

Prioritizing schizophrenia endophenotypes for future genetic studies: An example using data from the COGS-1 family study.

Journal

Schizophrenia research, 174(1-3)

ISSN

0920-9964

Authors

Millard, Steven P
Shofer, Jane
Braff, David
et al.

Publication Date

2016-07-01

DOI

10.1016/j.schres.2016.04.011

Peer reviewed

Schizophrenia Research 174 (2016) 1–9

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Prioritizing schizophrenia endophenotypes for future genetic studies: An example using data from the COGS-1 family study

Steven P. Millard [a,b], Jane Shofer [a,b], David Braff [c,d], Monica Calkins [e], Kristin Cadenhead [d], Robert Freedman [f], Michael F. Green [g,h], Tiffany A. Greenwood [d], Raquel Gur [e], Ruben Gur [e], Laura C. Lazzeroni [i], Gregory A. Light [c,d], Ann Olincy [f], Keith Nuechterlein [g], Larry Seidman [j], Larry Siever [k,l], Jeremy Silverman [k,l], William S. Stone [j], Joyce Sprock [c,d], Catherine A. Sugar [g,h,m], Neal R. Swerdlow [d], Ming Tsuang [d], Bruce Turetsky [e], Allen Radant [b,n], Debby W. Tsuang [b,n,*]

[a] VISN-20 Mental Illness Research, Education, and Clinical Center, VA Puget Sound Health Care System, Seattle, WA, USA
[b] VISN-20 Geriatric Research, Education, and Clinical Center, VA Puget Sound Health Care System, Seattle, WA, USA
[c] VISN-22 Mental Illness Research, Education, and Clinical Center, VA San Diego Healthcare System, San Diego, CA, USA
[d] Department of Psychiatry, University of California, San Diego, San Diego, CA, USA
[e] Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA
[f] Department of Psychiatry, University of Colorado Health Sciences Center, Denver, CO, USA
[g] Department of Psychiatry and Biobehavioral Sciences, Geffen School of Medicine at University of California, Los Angeles, Los Angeles, CA, USA
[h] VA Greater Los Angeles Health Care System, Los Angeles, CA, USA
[i] Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA, USA
[j] Massachusetts Mental Health Center Public Psychiatry Division of the Beth Israel Deaconess Medical Center, Harvard Medical School Department of Psychiatry, Boston, MA, USA
[k] Department of Psychiatry, Mount Sinai School of Medicine, New York, NY, USA
[l] VISN-3 Mental Illness Research, Education, and Clinical Center, James J. Peters VA Medical Center, New York, NY, USA
[m] Department of Biostatistics, Fielding School of Public Health at University of California, Los Angeles, Los Angeles, CA, USA
[n] Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA

ARTICLE INFO

Article history:
Received 5 January 2016
Received in revised form 6 April 2016
Accepted 11 April 2016
Available online 28 April 2016

Keywords:
Schizophrenia
Endophenotype
Logistic regression
Random forest
Sensitivity
Specificity
Accuracy
ROC curve
Multiple imputation

ABSTRACT

Past studies describe numerous endophenotypes associated with schizophrenia (SZ), but many endophenotypes may overlap in information they provide, and few studies have investigated the utility of a multivariate index to improve discrimination between SZ and healthy community comparison subjects (CCS). We investigated 16 endophenotypes from the first phase of the Consortium on the Genetics of Schizophrenia, a large, multi-site family study, to determine whether a subset could distinguish SZ probands and CCS just as well as using all 16.

Participants included 345 SZ probands and 517 CCS with a valid measure for at least one endophenotype. We used both logistic regression and random forest models to choose a subset of endophenotypes, adjusting for age, gender, smoking status, site, parent education, and the reading subtest of the Wide Range Achievement Test. As a sensitivity analysis, we re-fit models using multiple imputations to determine the effect of missing values.

We identified four important endophenotypes: antisaccade, Continuous Performance Test-Identical Pairs 3-digit version, California Verbal Learning Test, and emotion identification. The logistic regression model that used just these four endophenotypes produced essentially the same results as the model that used all 16 (84% vs. 85% accuracy).

While a subset of endophenotypes cannot replace clinical diagnosis nor encompass the complexity of the disease, it can aid in the design of future endophenotypic and genetic studies by reducing study cost and subject burden, simplifying sample enrichment, and improving the statistical power of locating those genetic regions associated with schizophrenia that may be the easiest to identify initially.

Published by Elsevier B.V.

1. Introduction

Schizophrenia (SZ) is a highly heritable yet complex multifactorial psychiatric disorder (Braff, 2015; Braff et al., 2007b; Thibaut, 2006). Locating genes that are associated with schizophrenia is a key step in identifying potentially remediable biological pathways for the development

* Corresponding author at: VA Puget Sound Health Care System, 1660 S Columbian Way, GRECC S182, Seattle, WA 98108, USA.
E-mail address: dwt1@uw.edu (D.W. Tsuang).

http://dx.doi.org/10.1016/j.schres.2016.04.011
0920-9964/Published by Elsevier B.V.

of novel treatments. Endophenotypes (e.g., neurocognitive and neurophysiologic measures) reflect components of liability narrower than the broad clinical diagnosis of schizophrenia and may facilitate the search for susceptibility genes and biological pathways to illness (Braff, 2015; Braff et al., 2007b; Braff et al., 2007c; Gottesman and Gould, 2003). Criteria for endophenotypes include: deficits associated with illness (moderate to large effect sizes between schizophrenia patients and community controls), state independence, heritability, and deficits in unaffected relatives at a higher rate than in the general population (small to moderate effect sizes between biological relatives of schizophrenia patients and community controls; Gur et al., 2007a). Iacono (1998) suggested an additional criterion for a useful endophenotype is that a deficit be unique to a class of related disorders (i.e., display specificity), but notes that "because psychiatric diagnosis is not perfectly reliable and the validity of the Diagnostic and Statistical Manual (DSM) diagnostic criteria is not firmly established, it is not reasonable to expect a complete absence of the endophenotype in an unrelated disorder that shows appreciable symptom overlap with the target disorder. Also, some endophenotypes will possibly index a dimension or process representing a dysfunction shared across certain disorders." Arfken et al. (2009) and Thibaut et al. (2015) distinguish between endophenotypes and biomarkers of a disease, such that biomarkers are disease specific, can be state- or trait-dependent, and are not necessarily heritable.

Because individual endophenotypes are more proximal functions of gene action than is the diagnostic assignment of schizophrenia itself, and individual endophenotypes are believed to reflect variation among a smaller number of genes than the very large array of genes implicated in schizophrenia, it should be simpler to localize the genetic loci contributing to the endophenotypes than to localize those for schizophrenia (Braff et al., 2007b; Braff et al., 2007c; Gottesman and Gould, 2003; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Iacono (1998) suggested that a multivariate phenotype might "have the greatest likelihood of assisting in the search for schizophrenia-related genes." While a number of studies have examined the association between individual endophenotypes and diagnosis, only a few have addressed the additive impact of using multiple endophenotypes simultaneously to improve discrimination between target populations. Comparing univariate and multivariate logistic regression, Price et al. (2006) showed that multivariate logistic regression using four neurophysiological endophenotypes (mismatch negativity, P50 suppression, P300 amplitude, and antisaccade error rate) to differentiate 60 SZ and 44 control subjects produced superior values of sensitivity (82%), specificity (73%), and accuracy (78%) compared to using just one endophenotype. They suggest that this multivariate endophenotype (i.e., the weighted linear combination using the coefficients from the multivariate logistic regression) could be used to increase power in genetic linkage and association analyses. Johannesen et al. (2013) investigated which subset of 14 neurophysiological endophenotypes best separated 50 SZ subjects from 50 healthy normal controls, and how that set of endophenotypes compared to the set that best separated the 50 SZ patients from 50 bipolar patients. The final model for separating SZ subjects from controls included five endophenotypes (P300 amplitude and latency, N100 Target, LFR Target, and GBR Standard), and yielded a sensitivity of 78%, specificity of 80%, and overall accuracy of 79%, but applying this model to distinguish SZ from bipolar patients yielded a sensitivity of 70%, specificity of 58%, and accuracy of 64%. The model that best separated SZ from bipolar patients included three endophenotypes (N100 S1, N100 Target, and GBR Target), and yielded a sensitivity of 74%, specificity of 70%, and overall accuracy of 72%. Peters et al. (2014) investigated the ability of 20 EEG indices to distinguish between 34 SZ and 37 control subjects. They first screened individual endophenotypes by testing for the significance of the area under the curve (AUC; equivalent to performing the nonparametric Mann-Whitney *U* test; Hanley and McNeil, 1982; Zweig and Campbell, 1993). Using the 11 indices that were significant, they performed principal components analysis (PCA), resulting in three top factors that explained 77% of the total variance. Using these three factors in a multivariate logistic regression model yielded 82% sensitivity, 70% specificity, and 76% accuracy; however, only one of the factors was a significant predictor in the model.

Numerous candidate endophenotypes for SZ have been proposed, yet there is substantial overlap in endophenotype performance between SZ probands and healthy controls. Furthermore, many endophenotypes may overlap in the information they provide. Identifying a subset of endophenotypes that taken together provide substantial sensitivity and specificity to differentiate between SZ and healthy comparison subjects has a two-fold research utility: practically, the degree to which the test battery can be limited is relevant to study cost, subject burden, and other design considerations, and furthermore, identification of key endophenotypes can help prioritize future genetic studies and improve statistical power by limiting the number of individual endophenotype statistical comparisons. We emphasize that we are not suggesting the use of a subset of endophenotypes to replace clinical diagnosis, nor are we suggesting that endophenotypes that are not part of the identified subset do not matter; rather, we are motivated by design efficiency in order to increase the power of locating those genetic regions associated with schizophrenia that may be the easiest to identify initially.

The goal of the first phase of the Consortium on the Genetics of Schizophrenia (COGS-1) was to investigate the genetic basis of endophenotypes for SZ (Calkins et al., 2007; Gur et al., 2007a). Results from COGS-1 have been reported by numerous authors (Greenwood et al., 2007; Greenwood et al., 2011; Greenwood et al., 2013; Horan et al., 2008; Light et al., 2014; Olincy et al., 2010; Radant et al., 2010; Radant et al., 2007; Radant et al., 2015; Stone et al., 2011; Swerdlow et al., 2007; Turetsky et al., 2007; Turetsky et al., 2008) and focus on describing the heritability and deficits in SZ probands of individual endophenotypes. The study described here uses statistical methods to investigate which combination of COGS-1 endophenotypes that we considered best distinguishes between SZ and healthy community comparison subjects (CCS) in a large, well-characterized sample.

2. Methods

2.1. Subjects

The current study utilizes data from the COGS-1 (Braff et al., 2007a; Braff et al., 2007b; Calkins et al., 2007), a multisite, family-based study on the genetics and heritability of neurocognitive and neurophysiologic SZ endophenotypes. Participants in this study met COGS-1 criteria as SZ probands (n = 345) or CCS (n = 517) and completed at least one endophenotype measure. The SZ subjects completed a structured clinical diagnostic interview (the Diagnostic Interview for Genetic Studies; Nurnberger et al., 1994) and a best-estimate consensus diagnostic procedure that included a comprehensive evaluation of psychosis, mood, and substance use disorders and related symptomology (Calkins et al., 2007; Greenwood et al., 2007). This information was used in concert with a standard medical record review to assess age, smoking status, education level, parental education level, handedness, and age at onset of psychosis; the Family Interview for Genetics Studies (FIGS; Maxwell, 1992) was used to assess subjects' family history of SZ or schizoaffective disorder; and the reading subtest of the Wide Range Achievement Test, 3rd edition (WRAT-3; Jastak and Wilkinson, 1993) was used to estimate premorbid intellectual functioning. The SZ subjects were also assessed using the Schedule for the Assessment of Negative Symptoms (SANS; Andreasen, 1983) and Schedule for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984); all probands met Diagnostic and Statistics Manual of Mental Disorders, 4th edition (DSM-IV), diagnostic criteria for SZ. Recruitment and ascertainment strategies, in-depth inclusion and exclusion criteria, in-person training and quality assurance guidelines for the assessment team and endophenotype testers,

informed consent procedures, and institutional review board approval at all seven study sites are detailed in Calkins et al. (2007). Inclusion criteria for SZ probands required that both biological parents were available for genotyping, and at least one full sibling unaffected by SZ was available for endophenotyping and genotyping.

2.2. Endophenotype assessment battery

The COGS-1 endophenotype assessment battery and its use in the different COGS-1 sub-studies have been discussed in detail elsewhere (see Introduction). We looked at one neurophysiological endophenotype, antisaccade performance (AS; Radant et al., 2007) measured as the ratio of correct antisaccades to total interpretable antisaccades, and 15 neurocognitive endophenotypes: the Degraded Stimulus (DS), 3-digit Identical Pairs (IP), and 4-digit IP versions of the Continuous Performance Test (CPT), which were measured using the signal/noise discrimination index (d′; Gur et al., 2007a); the Forward and Reordered condition of the Letter-Number Span (LNS; Horan et al., 2008) measured as the total number of correctly recalled sequences; the California Verbal Learning Test (CVLT; second edition), specifically the total and total semantic clustering scores on trials 1–5 (CVLT total and CVLT semantic; Stone et al., 2011); and the sensorimotor speed, motor speed, abstraction and mental flexibility, face memory, spatial memory, spatial ability, working memory, and emotion identification cognitive domains of the Penn Computerized Neurocognitive Battery (CNB; Gur et al., 2007a; Gur et al., 2007b). For the Penn CNB endophenotypes, all reported scores involving time were multiplied by −1 so that larger scores indicated better performance. Sensorimotor speed and motor speed were reported as z-scores based on the mean and standard deviation for the CCS, and for all other Penn CNB endophenotypes we used the efficiency scores, which were derived by averaging the z-scores for accuracy and speed for that particular measure. All of these endophenotypes were expected to show larger mean values for CCS compared to SZ subjects.

The COGS-1 study also included three other (neurophysiological) endophenotypes: prepulse inhibition (PPI; Swerdlow et al., 2007) and P50 suppression (measured as a ratio and a difference; Olincy et al., 2010). However, because of the large percentage of missing values for these endophenotypes (34% and 40%, respectively), we did not include them in our analyses.

2.3. Data analysis

Demographic differences between SZ and CCS were tested using Student's *t*-test for continuous variables and the chi-squared test for categorical variables. For categorical variables, confidence intervals for the difference between two proportions were based on inverting the score statistic (Newcombe, 1998).

As described in Price et al. (2006) and Johannesen et al. (2013), we first compared unadjusted endophenotypes between groups using summary statistics, strip charts, and two-sample *t*-tests (effect size was computed using Cohen's *d* statistic). We also computed distribution by group of the number of endophenotypes in the first quartile (based on combining groups), and the correlations between endophenotypes. Next, we used both logistic regression and random forest models (Breiman, 2001; Kuhn and Johnson, 2013; Liaw and Wiener, 2002) to determine which set of endophenotypes best distinguished between SZ probands and CCS. Following previous COGS-1 studies comparing endophenotype performance in SZ versus CCS, all models included the covariates age, gender, smoking status, site, maximum of parents' education, and WRAT-3 score. The logistic regression and random forest models included only subjects with non-missing values for the model covariates and all 16 endophenotypes.

For logistic regression, the first model included only the covariates, denoted *LR 0 EPs*. The second model entered each endophenotype by itself (i.e., 16 separate models), denoted *LR 1 EP*. The third model included all 16 endophenotypes simultaneously as predictor variables, denoted *LR 16 EPs*. The fourth model, denoted *LR Top EPs*, used the smallest subset of endophenotypes that were indicated as important based on the results from forward and backward stepwise logistic regression (based on the Bayesian Information Criterion [BIC]) and random forest [see below]). The association between endophenotype and proband status was summarized with the odds ratios (ORs) for proband, given a decrease (deficit) in endophenotype score equivalent to its interquartile range (Harrell, 2001). Confidence intervals for the ORs were based on Wald estimates.

Logistic regression models require explicit modeling of nonlinear and interaction terms, and stepwise methods may depend on which criterion is used for inclusion or deletion. Random forest models, however, are nonparametric, inherently account for nonlinearity and interactions, and have built-in cross-validation (Breiman, 2001; Kuhn and Johnson, 2013; Liaw and Wiener, 2002). We therefore also ran a random forest model using the 16 endophenotypes (denoted *RF 16 EPs*); variable importance was computed based on the unscaled mean decrease in accuracy (Nicodemus et al., 2010).

Models were compared using three metrics: generalized R^2 values applicable to logistic regression (roughly the proportion of the variability in the response, i.e., SZ vs CCS, attributable to the estimated model; Harrell, 2001); sensitivity, specificity, and accuracy with confidence intervals (sensitivity and specificity presented with joint confidence intervals; Pepe, 2003); and Somer's D_{xy} rank correlation (Harrell, 2001) between predicted probabilities and observed responses, which measures how well the model discriminates between SZ and CCS ($D_{xy} = 1$ denotes perfect discrimination, whereas $D_{xy} = 0$ implies that the model does no better than a random assignment into groups; D_{xy} is equivalent to the area under the curve: $D_{xy} = 2(AUC - 0.5)$). For the logistic regression models, both the generalized R^2 and Somer's D_{xy} are presented unadjusted, as well as corrected for optimism (i.e., model overfitting; Harrell, 2001) based on resampling. ROC curves were calculated to provide a visual comparison of models.

Because of the large proportion of subjects with missing values for at least one of the endophenotypes, as a sensitivity analysis we repeated the logistic regressions and random forest using multiple imputations (Buuren and Groothuis-Oudshoorn, 2011). The imputations were performed using the R package *mice* (Multiple Imputation by Chained Equations), using predictive mean matching. All 16 endophenotypes and the model covariates were used in the imputations. Fifty imputed datasets were estimated. Pooled summary statistics were obtained by averaging the estimates of models across the 50 imputed datasets. Pooled total variances for summary statistics were estimated using Rubin's formula (3.1.5; Rubin, 1987).

All analyses were performed using R version 3.2.0 (R Development Core Team, 2011), the *EnvStats* package (Millard, 2013) to compute summary statistics and plot strip charts, the *randomForest* package (Liaw and Wiener, 2002), the *rms* package (Harrell, 2013) to carry out logistic regressions and compute the R^2 and Somer's D_{xy} values, the *pROC* package (Robin et al., 2011) to construct the ROC curves, and the *mice* package (Buuren and Groothuis-Oudshoorn, 2011) to perform the multiple imputations.

3. Results

3.1. Demographic and clinical variables

We obtained data on endophenotypes from 862 subjects: 345 SZ and 517 CCS. SZ probands had an average age of onset of 21 years and were more likely to be male, white, and smokers and to have less education and a lower WRAT-3 score compared to CCS; however, the parents of SZ probands were slightly more educated compared to the parents of CCS (Table 1). There were 802 subjects (318 SZ, 484 CCS) with non-missing data for the model covariates age, gender, smoking status, site, maximum of parents' education, and WRAT-3 score. Supplemental

Table 1
Demographic and clinical characteristics of schizophrenia (SZ) and community comparison subjects (CCS).

	Community comparison subjects (CCS) (n = 517)[a]	Schizophrenia subjects (SZ) (n = 345)[a]	*P*-value[b]	Difference (CCS – SZ) (95% CI)[b]
Age (years)	36.3 (12.7) [18, 65]	35.0 (11.1) [18, 62]	0.12	1.3 [−0.4, 2.9]
Gender (male)	223 (43%)	262 (76%)	<0.0001	−33% [−39%, −26%]
Race (white)	345 (67%)	254 (74%)	0.04	−7% [−13%, −0.5%]
Smoker[c]	70 (14%)	158 (46%)	<0.0001	−32% [−39%, −27%]
Education (years)[d]	15.3 (2.4) [8, 22]	13.5 (2.1) [8, 20]	<0.0001	1.8 [1.6, 2.2]
Max parents' education (years)[e]	15.0 (3.2) [2, 25]	15.6 (3.5) [0, 25]	0.008	−0.6 [−1.1, −0.2]
WRAT-3 score[f]	107.2 (10.6) [70, 125]	101.8 (11.7) [64, 122]	<0.0001	5.4 [3.8, 6.9]
Age at onset of symptoms (years)[g]		20.9 (5.5) [6, 51]		
SANS[h]		9.6 (5.9) [0, 25]		
SAPS[i]		6.3 (4.1) [0, 20]		

Abbreviations: CCS, community comparison subjects; CI, confidence interval; SANS, Schedule for the Assessment of Negative Symptoms; SAPS, Schedule for the Assessment of Positive Symptoms; SD, standard deviation; SZ, schizophrenia subjects; WRAT-3, Wide Range Achievement Test, 3rd edition.

[a] Mean (SD) [Min, Max] presented for continuous variables. Number (%) presented for categorical variables.

[b] Between-group (CCS – SZ) *P*-values and confidence intervals are based on the chi-squared test for categorical variables and the *t*-test for continuous variables.

[c] 1 missing value (0.2%) for CCS, 5 missing values (1%) for SZ.

[d] 1 missing value (0.2%) for CCS, 2 missing values (1%) for SZ.

[e] 29 missing values (6%) for CCS, 9 missing values (3%) for SZ.

[f] 4 missing values (1%) for CCS, 15 missing values (4%) for SZ.

[g] 8 missing values (2%) for SZ.

[h] 9 missing value (3%) for SZ.

[i] 9 missing value (3%) for SZ.

Fig. S1 presents a flow chart showing sample sizes by presence of endophenotypes and model covariates, and Supplemental Table S1 presents the demographic and clinical characteristics of the 571 subjects (174 SZ and 397 CCS) with complete data for all 16 endophenotypes and model covariates.

3.2. Discriminatory effectiveness of endophenotypes

Table 2 presents descriptive statistics for the endophenotypes by subject group, as well as the results of two-sample *t*-tests to compare groups. As expected, all endophenotypes showed larger mean values for CCS compared to SZ subjects. SZ subjects had higher rates of missing values compared to CCS for all endophenotypes. Results based on including only the 571 subjects with complete data for all 16 endophenotypes and adjusting for model covariates were similar (Supplemental Table S2), except that DS-CPT was no longer significant. Supplemental Fig. S2 uses strip charts with 95% confidence intervals to illustrate the distributions of endophenotypes by group without adjusting for covariates. The extremely large overlap in performance between groups is evident for all endophenotypes. However, Supplemental Table S3, which shows the distribution by group of the number of endophenotypes in the first quartile (based on both unadjusted and adjusted scores), indicates that schizophrenia probands are more likely than community control subjects to have lower scores on more endophenotypes. Supplemental Table S4 displays pairwise endophenotype correlations. Emotion identification is relatively highly correlated with all other Penn CNB endophenotypes, except for motor speed (which is only moderately correlated with any of the other Penn CNB endophenotypes).

Table 3 displays the results of the logistic regression models used to discriminate between SZ probands and CCS based on various combinations of the 16 endophenotypes. The displayed ORs are based on comparing the odds of declaring a subject a proband assuming the endophenotype is equal to its 25th percentile versus the same odds assuming the endophenotype is equal to its 75th percentile. For example, looking at the results for the antisaccade task using model *LR 1 EP*, the odds of declaring a subject SZ given a score of 0.64 (the 25th percentile) divided by the same odds given a score of 0.91 (the 75th percentile) was 4.09 with a 95% confidence interval of [2.86, 5.83]. For model *LR 16 EPs*, in which all 16 endophenotypes were included in the same model, five of the endophenotypes had ORs significantly >1, and two (DS-CPT and spatial ability) unexpectedly had ORs <1; this result is likely due to collinearity among the endophenotypes, in particular, spatial ability had the highest variable inflation factor (2.4) of the 16 endophenotypes. The forward stepwise regression model included the endophenotypes antisaccade, CPT-IP 3-digit, CVLT total, emotion identification, and motor speed, while the backward stepwise regression model included these same endophenotypes as well as sensorimotor speed and spatial ability. However, the random forest model (model *RF 16 EPs*; Supplemental Fig. S3) clearly indicates that the most important endophenotypes are just the first four (i.e., antisaccade, CPT-IP 3-digit, CVLT total, and emotion identification; model *LR Top EPs*). There is very little difference in the generalized R^2 or Somer's D_{xy} values between the model that uses all 16 endophenotypes (*LR 16 EPs*) and the one that uses just the top four (*LR Top EPs*), and both of these models are superior to using the model that includes just the covariates, model *LR 0 EPs*, which had R^2 and Somer's D_{xy} values of 0.32 and 0.59, respectively. Supplemental Fig. S4 illustrates Price et al.'s (2006) multivariate endophenotype concept: strip charts by group of the predicted probability of being assigned to the SZ group based on the *LR Top EPs* model.

Table 4 shows the sensitivity, specificity, and accuracy of models *LR 16 EPs*, *LR Top EPs*, and *RF 16 EPs*, and Fig. 1 shows the associated ROC curves. Not surprisingly, the ROC curve and Somer's D_{xy} based on the random forest are lower compared to those based on the logistic regression models because random forest has built-in cross-validation whereas standard logistic regression models do not. Using a predicted probability of 0.5 as the cut-off for declaring a subject to be SZ, the sensitivity is much smaller than the specificity, whereas using a cut-off of 0.3 makes the sensitivity approximately equal to the specificity.

As a sensitivity analysis, the logistic regression and random forest models were re-estimated based on multiple imputation of missing values using fifty imputations. Supplemental Fig. S5 shows that the ORs and 95% confidence intervals for model *LR Top EPs* using the original data and based on the multiple imputations are similar. Supplemental Table S5 compares sensitivity, specificity, and accuracy based on the original versus imputed data. Point estimates for specificity and accuracy were similar between the imputation and original analyses. However, point estimates for sensitivity based on the imputation analysis were about 10 and 15 percentage points higher, respectively, for the logistic regression and random forest models compared to the original analysis. This is because for almost all endophenotypes, the mean endophenotype value for the imputed values was less than the mean for the non-missing data, and since more SZ than CCS subjects had

Table 2
Unadjusted endophenotypes for schizophrenia (SZ) and community comparison subjects (CCS). All endophenotypes were expected to show larger mean values for CCS compared to SZ.

	Community comparison subjects (CCS) (n = 517)		Schizophrenia subjects (SZ) (n = 345)		*P*-value[a]	Effect size[a]	Difference (CCS – SZ)	95% CI[a]
	Mean (SD) [range]	Number missing (%)	Mean (SD) [range]	Number missing (%)				
Antisaccade (proportion correct)[b]	0.82 (0.15) [0.04, 1.00]	22 (4%)	0.61 (0.26) [0.00, 0.98]	61 (18%)	<0.0001	1.07	0.21	(0.18, 0.24)
DS-CPT (d′)[c]	2.54 (1.03) [−0.12, 5.42]	17 (3%)	2.35 (1.14) [0.05, 5.42]	63 (18%)	0.014	0.18	0.20	(0.04, 0.35)
CPT-IP 3-digit (d′)[d]	2.99 (0.80) [0.92, 4.79]	23 (4%)	2.15 (0.84) [−0.25, 4.52]	70 (20%)	<0.0001	1.03	0.84	(0.72, 0.96)
CPT-IP 4-digit (d′)[e]	1.96 (0.84) [0.05, 4.26]	23 (4%)	1.28 (0.71) [−0.17, 3.51]	71 (21%)	<0.0001	0.85	0.68	(0.56, 0.80)
LNS forward (number correct)[f]	14.3 (2.9) [7, 21]	3 (1%)	12.9 (2.9) [4, 21]	13 (4%)	<0.0001	0.48	1.4	(1.0, 1.8)
LNS reordered (number correct)[g]	11.4 (2.7) [2, 20]	2 (<1%)	9.1 (2.9) [1, 17]	19 (6%)	<0.0001	0.81	2.3	(1.9, 2.7)
CVLT total (number correct)[h]	56.2 (10.3) [21, 80]	5 (1%)	41.6 (12.8) [6, 75]	24 (7%)	<0.0001	1.29	14.6	(13.1, 16.2)
CVLT semantic (number correct)[i]	2.0 (2.4) [−3.0, 8.5]	5 (1%)	0.4 (1.5) [−2.2, 7.6]	16 (5%)	<0.0001	0.79	1.6	(1.4, 1.9)
Penn CNB								
Speed[j]								
Sensorimotor	0.00 (1.00) [−12.1, 1.3]	25 (5%)	−0.80 (1.88) [−12.8, 1.3]	41 (12%)	<0.0001	0.58	0.80	(0.61, 1.00)
Motor	0.00 (1.00) [−4.9, 4.6]	27 (5%)	−0.70 (1.43) [−5.9, 4.5]	43 (12%)	<0.0001	0.59	0.70	(0.53, 0.87)
Efficiency (mean of 2 z-scores)[k]								
Abstraction/mental flexibility	0.00 (0.85) [−4.5, 1.1]	32 (6%)	−0.63 (1.15) [−6.7, 0.9]	49 (14%)	<0.0001	0.65	0.63	(0.49, 0.77)
Face memory	0.01 (0.84) [−10.3, 1.4]	23 (4%)	−0.71 (1.21) [−10.4, 1.0]	37 (11%)	<0.0001	0.70	0.70	(0.56, 0.84)
Spatial memory	0.00 (0.64) [−3.2, 1.3]	35 (7%)	−0.52 (0.98) [−6.0, 1.4]	54 (16%)	<0.0001	0.66	0.52	(0.41, 0.64)
Spatial ability	0.00 (0.78) [−4.3, 1.2]	33 (6%)	−0.27 (1.01) [−6.3, 1.2]	60 (17%)	<0.0001	0.31	0.27	(0.14, 0.40)
Working memory	0.00 (0.77) [−3.5, 1.1]	29 (6%)	−0.63 (1.13) [−5.2, 1.1]	68 (20%)	<0.0001	0.68	0.63	(0.49, 0.76)
Emotion identification	0.00 (0.78) [−4.9, 1.5]	25 (5%)	−1.19 (1.45) [−7.9, 1.0]	42 (12%)	<0.0001	1.10	1.19	(1.03, 1.34)

Abbreviations: CCS, community comparison subjects; CI, confidence interval; SD, standard deviation; CNB, Computerized Neurocognitive Battery; CPT-IP, Continuous Performance Test, Identical Pairs version; CVLT, California Verbal Learning Test; DS-CPT, Degraded Stimulus Continuous Performance Test; LNS, Letter-Number Span; SZ, schizophrenia subjects.

[a] *P*-value and 95% CI for difference in means based on two-sample *t*-test assuming equal variances. Effect size computed using Cohen's *d* statistic.

[b] Proportion correct out of a maximum of 60 trials.

[c] Overall signal/noise discrimination (d′).

[d] Three-digit d′.

[e] Four-digit d′.

[f] After each sequence, the participant is asked to recall the numbers and letters in the same exact order, with no reordering of the stimuli. The number of digits and letters increases by one on each trial, from one up to a maximum length of 8 stimuli. Three sequences of the same length are presented during each trial. The test is discontinued when the subject fails three consecutive sequences of the same length. The score is the total number of correctly recalled sequences.

[g] After each sequence, the participant is asked to repeat the numbers in ascending order first and then the letters in alphabetical order.

[h] Trials 1–5 Free Recall Correct.

[i] Total semantic clustering scores on trials 1–5.

[j] Sensorimotor and motor values were reported as z-scores based on the mean and standard deviation for the community comparison subjects. Z-scores were based on values that had been multiplied by −1, so that a larger value indicated a better performance.

[k] Efficiency scores were derived by averaging the z-scores (based on the mean and standard deviation for the community comparison subjects) for accuracy and speed.

missing data, more of the imputed low scoring outcomes were also from the SZ group, increasing the probability that a low scoring subject was from the SZ group.

4. Discussion

The endophenotype concept is a tool that may potentially identify key genetic loci associated with schizophrenia. However, the enormous overlap in endophenotype performance between schizophrenia and healthy community control subjects displayed in the strip charts shown in Supplemental Fig. S2 emphasizes the complex nature of the disease. We have shown that for the COGS-1 study, using just four endophenotypes (antisaccade, CPT-IP 3-digit, CVLT total, and emotion identification) produces essentially the same discrimination between SZ probands and CCS as using all 16 endophenotypes that we considered. This finding is consistent both between models (logistic regression versus random forest) and datasets (available versus imputed). The heritabilities [95% CI] of these endophenotypes have been reported as 0.42 [0.27, 0.57], 0.38 [0.23, 0.52], 0.25 [0.11, 0.40], and 0.32 [0.18, 0.46], respectively (Greenwood et al., 2007), and are similar to the estimated heritability of SZ itself of 31% and 44% for nuclear and extended families, respectively, based on the COGS-1 data (Light et al., 2014). Greenwood et al. (2013) performed a genome-wide linkage analysis on the COGS-1 data, and although the study had limited power, the antisaccade task showed significant genome-wide linkage (LOD of 4 for chromosome 3p14) and emotion recognition almost attained significant genome-wide linkage (LOD of 3.5 for chromosome 1p36). The four key endophenotypes we identified also have the largest effect sizes (Table 2 and Supplemental Table S2); however, in general, endophenotypes with smaller effect sizes may end up being identified as important for separating groups using the methods presented here,

Table 3
Results of logistic regression analyses using subjects with complete data for 16 endophenotypes. Odds ratios (95% CI based on Wald estimate) for schizophrenia subjects (SZ) versus CCS for a decrease (deficit) in endophenotype equivalent to the endophenotype interquartile range. ORs in bold denote a significant difference from 1 ($P < 0.05$).[a*]

	Model (n = 571)*		
	LR 1 EP	*LR 16 EPs*	*LR Top EPs*
Antisaccade	**4.09 (2.86, 5.83)**	**3.45 (2.21, 5.39)**	**2.93 (1.97, 4.37)**
DS-CPT	1.06 (0.81, 1.39)	**0.68 (0.46, 0.99)**	
CPT-IP 3-digit	**4.98 (3.23, 7.67)**	1.67 (0.88, 3.16)	**2.26 (1.38, 3.71)**
CPT-IP 4-digit	**3.66 (2.47, 5.43)**	1.80 (0.98, 3.31)	
LNS forward	**1.41 (1.02, 1.95)**	0.93 (0.58, 1.48)	
LNS reordered	**1.75 (1.33, 2.32)**	0.96 (0.62, 1.47)	
CVLT total	**6.75 (4.10, 11.1)**	**2.18 (1.04, 4.59)**	**3.51 (1.99, 6.18)**
CVLT semantic	**2.22 (1.64, 3.01)**	1.50 (0.97, 2.34)	
Penn CNB			
Sensorimotor	**1.78 (1.38, 2.29)**	**1.59 (1.12, 2.26)**	
Motor	**2.01 (1.50, 2.69)**	**1.60 (1.11, 2.30)**	
Abstraction/mental flexibility	**2.64 (1.79, 3.88)**	1.14 (0.64, 2.01)	
Face memory	**2.45 (1.74, 3.45)**	1.37 (0.84, 2.23)	
Spatial memory	**1.75 (1.30, 2.36)**	0.80 (0.51, 1.24)	
Spatial ability	**1.54 (1.14, 2.07)**	**0.49 (0.29, 0.82)**	
Working memory	**2.32 (1.67, 3.22)**	1.07 (0.67, 1.70)	
Emotion identification	**3.43 (2.40, 4.90)**	**2.36 (1.51, 3.68)**	**2.41 (1.65, 3.53)**
R^2 (corrected for optimism)		0.63 (0.56)	0.59 (0.54)
Somer's D_{xy} (corrected for optimism)		0.85 (0.80)	0.81 (0.79)

Abbreviations: CCS, community comparison subjects; CNB, Computerized Neurocognitive Battery; CPT-IP, Continuous Performance Test, Identical Pairs version; CVLT, California Verbal Learning Test; DS-CPT, Degraded Stimulus Continuous Performance Test; LNS, Letter-Number Span; OR, odds ratio; SZ, schizophrenia subjects.

[a] All models were adjusted for age, gender, site, parents' education, WRAT-3 and smoking status. There are n = 802 subjects with all these covariates present, but only n = 571 subjects (n = 174 SZ, n = 397 CCS) with all covariates and all 16 endophenotypes present. See Supplemental Fig. S1 for a flow chart showing sample sizes. *LR 1 EP.* Univariate models (i.e., only one endophenotype included in the model at a time, along with covariates). *LR 16 EPs.* Multivariate model that includes all 16 endophenotypes, along with covariates. *LR Top EPs.* Multivariate model that includes the top four endophenotypes based on forward and backward stepwise regression and random forest, along with covariates.

because a variable can become important once other variables are in the model (Sun et al., 1996).

Identifying a subset of endophenotypes with good discriminatory characteristics can help prioritize future endophenotypic and genetic studies in schizophrenia. Limiting the test battery will reduce cost and subject burden, and limiting the number of individual endophenotype comparisons can increase statistical power. Studies that enrich the sample by choosing controls that perform well on all or most endophenotypes and/or probands that perform poorly on all or most endophenotypes (Supplemental Table S3b) will be easier to conduct, and genetic analyses could involve the multivariate endophenotype (Supplemental Fig. S4).

Our study is similar to that of Price et al. (2006) and Johannesen et al. (2013) except that we used a random forest model in addition to logistic regression models, and our results are based on one of the largest study samples to date that underwent uniform endophenotypic assessments. Our results for sensitivity (85%) and specificity (85%) based on the logistic regression model using just the four endophenotypes (model *LR Top EPs*) and a cut-off of 0.3 for classification are superior to those of Price et al. (2006), Johannesen et al. (2013), and Peters et al. (2014). However, more realistic estimates for future studies based on the random forest model indicate a sensitivity of 81% and specificity of 74%, which are similar to those of the three previous studies.

Some limitations of the previous studies should be discussed. First, in addition to sensitivity, specificity, and accuracy, Price et al. (2006) present estimates of positive predictive value (PPV) and negative predictive value (NPV); however, correct estimates of PPV and NPV require knowing the prevalence of the disease and therefore require using a cohort (not case-control) design or else an external estimate of the prevalence in the computation of the estimates (Pepe, 2003; Pepe et al., 2004; Zweig and Campbell, 1993). Second, given that the focus of these studies was on using logistic regression models, there is no constraint on the distribution of the predictor variables (i.e., the endophenotypes), yet Johannesen et al. (2013) removed "outliers" based on looking at boxplots. Third, Peters et al. (2014) selected variables for inclusion by using results from univariate tests, which is not an optimal method to determine the most important variables for a multivariate model, because a variable by itself may not significantly distinguish between groups but can become significant once other variables are in the model (Sun et al., 1996). Fourth, it is not clear whether Peters et al. (2014) performed PCA on the raw data or the normalized (i.e., z-transformed) data. If the variables are measured on scales with widely different ranges, then the results of the PCA based on the raw data will be dominated by the variables with the largest range (Johnson and Wichern, 2007).

Supplemental Table S3 shows that the four endophenotypes we identified are moderately correlated with each other ($r = 0.35$ to 0.45), except for the correlation between antisaccade and emotion identification ($r = 0.28$). However, these endophenotypes are associated with diverse neurological functions: the antisaccade task measures the ability of the oculomotor system to inhibit prepotent responses (Radant et al., 2007), the Continuous Performance Test-Identical Pairs

Table 4
Sensitivity, specificity, and accuracy with 95% confidence intervals for logistic regression and random forest models for subjects with non-missing values for all 16 endophenotypes (n = 571).

	Classify as proband when model probability > 0.5			Classify as proband when model probability > 0.3			
Model	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Somer's D_{xy} raw (adjusted)
LR 16 EPs	70 (62, 78)	91 (88, 95)	85 (82, 88)	88 (82, 94)	86 (82, 90)	86 (84, 89)	0.85 (0.80)
LR Top EPs	66 (58, 74)	92 (89, 95)	84 (81, 87)	85 (78, 91)	85 (81, 89)	85 (82, 88)	0.81 (0.79)
RF 16 EPs	56 (48, 65)	92 (90, 95)	76 (73, 80)	81 (74, 88)	74 (69, 79)	81 (78, 85)	0.72

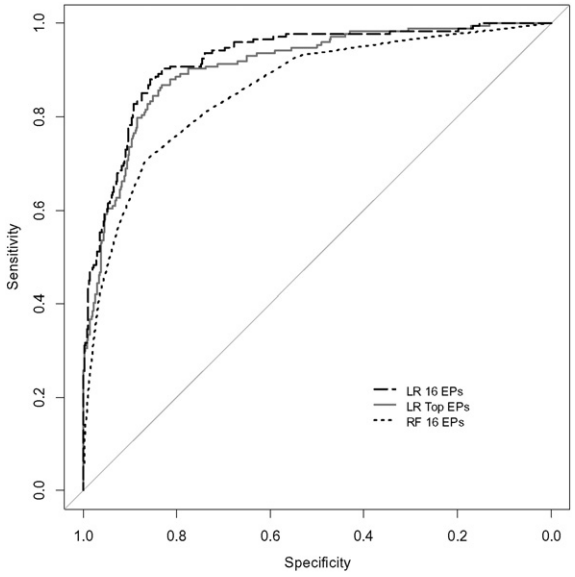

Fig. 1. ROC curves for logistic regression and random forest models for subjects with non-missing values for all 16 endophenotypes (n = 571).

numbers version measures sustained focused attention with a demand on working memory (Cornblatt and Keilp, 1994), the California Verbal Learning Test measures verbal declarative memory (Stone et al., 2011), and the Penn CNB emotion identification task measures social cognition (Gur et al., 2010).

Using 12 of the 16 COGS-1 endophenotypes presented here (CPT-IP 3-digit, CVLT semantic, LNS forward, and Penn CNB working memory were omitted), as well as PPI, P50 gating, and an additional DS-CPT measure (reaction time for targets), Seidman et al. (2015) used factor analysis applied to SZ probands (n = 83), CCS (n = 209), and SZ nonpsychotic siblings (n = 151) with non-missing values for all 15 endophenotypes to identify five factors, and then looked at endophenotype loading on each factor, the correlation between factors, and the heritability of the factors. Ten of the 15 endophenotypes were associated with at least one of the factors, and the factor structure was similar within each diagnostic subgroup. Based on the loadings, the five factors were identified as reflecting primarily episodic memory, working memory, perceptual vigilance, visual abstraction, and inhibitory processing, respectively, with heritabilities ranging from 22% to 39%. Unlike our study, which focuses on which subset of endophenotypes best distinguishes SZ probands from CCS, the study by Seidman et al. (2015) sought to identify underlying constructs (factors) that are responsible for the correlations between endophenotypes (Johnson and Wichern, 2007). The four endophenotypes identified in our study load on Seidman et al.'s episodic memory (antisaccade, CVLT total, and emotion recognition), working memory (CVLT total), and inhibitory process (antisaccade) factors (Seidman et al. did not include CPT-IP 3-digit, but CPT-IP 4-digit loads on working memory).

There are many strengths to our study. The data are based on one of the largest study samples to date that has undergone uniform endophenotypic assessments. Unlike past studies investigating multivariate endophenotypes that used only neurophysiological endophenotypes, we used both neurophysiological and neurocognitive endophenotypes. We used multiple methods to determine which subset of endophenotypes best distinguishes diagnostic groups, including forward stepwise regression, backward stepwise regression, and random forest. Finally, we used multiple imputations to show that our results are essentially the same after imputing missing values.

There are also limitations to our study. The number of neurophysiological endophenotypes available was limited compared to past studies. (Neurophysiological endophenotype collection in a substantially larger cohort of SZ cases and healthy comparison subjects from our COGS-2 study has been reported [Swerdlow et al., 2015], and genetic analyses are forthcoming.) We also did not have equivalent measurements in related diagnostic groups (e.g., bipolar disorder), which would be necessary to examine whether the combination of endophenotypes identified in this study will also show good discrimination between SZ and other psychiatric diagnostic groups. Finally, the COGS-1 inclusion requirement that both biological parents be available for genotyping, and at least one full sibling unaffected by SZ be available for endophenotyping and genotyping, may limit the generalizability of our results.

In conclusion, our study confirms past studies that showed that using a multivariate approach to endophenotype-based indices for discriminating SZ probands from healthy controls yields sensitivity and specificity on the order of 80% using just a few endophenotypes, and in our study, four endophenotypes produced essentially the same discrimination between groups as using all 16 endophenotypes. These findings, and applying the methods we used here to other datasets that may include endophenotypes we did not look at here, can help direct future studies by helping researchers decide which endophenotypes to include in a test battery and improving statistical power by limiting the number of individual endophenotype statistical comparisons.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.schres.2016.04.011.

Conflicts of interest

Dr. Green has been a consultant to AbbVie, DSP, Forum, Mnemosyne (scientific board), and Takeda, and received research support from Amgen and Forum. Dr. Lazzeroni is an inventor on a patent application filed by Stanford University on genetic polymorphisms associated with depression. Dr. Light has been a consultant to EnVivo/Forum and Astellas and serves on an advisory board for Neuroverse. Dr. Nuechterlein has received unrelated research support from Janssen Scientific Affairs, Genentech, and Brain Plasticity, Inc., and he has consulted for Genentech, Otsuka, Janssen, and Brain Plasticity, Inc. Dr. Swerdlow has been a consultant for Genco Sciences, Ltd. All other authors declare that they have no conflict of interest.

Contributors

Dr. Millard, Ms. Shofer, and Dr. Tsuang composed the manuscript. Dr. Millard and Ms. Shofer performed the statistical analyses. All other authors participated in aspects of the COGS-1 study design, including subject recruitment, endophenotype testing, and validation of the clinical and endophenotype data. All authors were responsible for reviewing, editing, and approving the final version of the manuscript.

Funding

This material is the result of work supported in part with resources and the use of facilities at the VA Puget Sound Health Care System, Seattle, WA; VA San Diego Healthcare System, San Diego, CA; VA Greater Los Angeles Health Care System, Los Angeles, CA; and James J. Peters VA Medical Center, New York, NY. The study was supported by NIMH grants R01 MH65571, R01 MH042228, R01 MH65588, R01 MH65562, R01 MH65707, R01 MH65554, R01 MH65578, R01 MH086135, and R01 MH65558. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript, and the contents do not represent the views of the funders, the U.S. Department of Veterans Affairs, or the United States Government.

Acknowledgments

The authors wish to thank all of the subjects and their family members for participating in the COGS-1 study. We also thank Andrew Shutes-David for his editorial assistance, all of the research staff for their work on this study, and the reviewers for their insightful and helpful comments.

References

Andreasen, N., 1983. The Scale for the Assessment of Negative Symptoms (SANS). University of Iowa, Iowa City, IA.
Andreasen, N., 1984. The Scale for the Assessment of Positive Symptoms (SAPS). University of Iowa, Iowa City, IA.
Arfken, C.L., Carney, S., Boutros, N.N., 2009. Translating biological parameters into clinically useful diagnostic tests. Curr. Psychiatry Rep. 11 (4), 320–323.
Braff, D.L., 2015. The importance of endophenotypes in schizophrenia research. Schizophr. Res. 163 (1–3), 1–8.
Braff, D., Schork, N.J., Gottesman, I.I., 2007a. Endophenotyping schizophrenia. Am. J. Psychiatry 164 (5), 705–707.

Braff, D.L., Freedman, R., Schork, N.J., Gottesman, I.I., 2007b. Deconstructing schizophrenia: an overview of the use of endophenotypes in order to understand a complex disorder. Schizophr. Bull. 33 (1), 21–32.
Braff, D.L., Light, G.A., Swerdlow, N.R., 2007c. Prepulse inhibition and P50 suppression are both deficient but not correlated in schizophrenia patients. Biol. Psychiatry 61 (10), 1204–1207.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Buuren, S., Groothuis-Oudshoorn, K., 2011. Mice: multivariate imputation by chained equations in R. J. Stat. Softw. 45 (3), 1–67.
Calkins, M.E., Dobie, D.J., Cadenhead, K.S., Olincy, A., Freedman, R., Green, M.F., Greenwood, T.A., Gur, R.E., Gur, R.C., Light, G.A., Mintz, J., Nuechterlein, K.H., Radant, A.D., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Braff, D.L., 2007. The consortium on the genetics of endophenotypes in schizophrenia: model recruitment, assessment, and endophenotyping methods for a multisite collaboration. Schizophr. Bull. 33 (1), 33–48.
Cornblatt, B.A., Keilp, J.G., 1994. Impaired attention, genetics, and the pathophysiology of schizophrenia. Schizophr. Bull. 20 (1), 31–46.
R Development Core Team, 2011. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Gottesman, I.I., Gould, T.D., 2003. The endophenotype concept in psychiatry: etymology and strategic intentions. Am. J. Psychiatry 160 (4), 636–645.
Greenwood, T.A., Braff, D.L., Light, G.A., Cadenhead, K.S., Calkins, M.E., Dobie, D.J., Freedman, R., Green, M.F., Gur, R.E., Gur, R.C., Mintz, J., Nuechterlein, K.H., Olincy, A., Radant, A.D., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Schork, N.J., 2007. Initial heritability analyses of endophenotypic measures for schizophrenia: the consortium on the genetics of schizophrenia. Arch. Gen. Psychiatry 64 (11), 1242–1250.
Greenwood, T.A., Lazzeroni, L.C., Murray, S.S., Cadenhead, K.S., Calkins, M.E., Dobie, D.J., Green, M.F., Gur, R.E., Gur, R.C., Hardiman, G., Kelsoe, J.R., Leonard, S., Light, G.A., Nuechterlein, K.H., Olincy, A., Radant, A.D., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Freedman, R., Braff, D.L., 2011. Analysis of 94 candidate genes and 12 endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia. Am. J. Psychiatry 168 (9), 930–946.
Greenwood, T.A., Swerdlow, N.R., Gur, R.E., Cadenhead, K.S., Calkins, M.E., Dobie, D.J., Freedman, R., Green, M.F., Gur, R.C., Lazzeroni, L.C., Nuechterlein, K.H., Olincy, A., Radant, A.D., Ray, A., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Sugar, C.A., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Light, G.A., Braff, D.L., 2013. Genome-wide linkage analyses of 12 endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia. Am. J. Psychiatry 170 (5), 521–532.
Gur, R.E., Calkins, M.E., Gur, R.C., Horan, W.P., Nuechterlein, K.H., Seidman, L.J., Stone, W.S., 2007a. The Consortium on the Genetics of Schizophrenia: neurocognitive endophenotypes. Schizophr. Bull. 33 (1), 49–68.
Gur, R.E., Nimgaonkar, V.L., Almasy, L., Calkins, M.E., Ragland, J.D., Pogue-Geile, M.F., Kanes, S., Blangero, J., Gur, R.C., 2007b. Neurocognitive endophenotypes in a multiplex multigenerational family study of schizophrenia. Am. J. Psychiatry 164 (5), 813–819.
Gur, R.C., Richard, J., Hughett, P., Calkins, M.E., Macy, L., Bilker, W.B., Brensinger, C., Gur, R.E., 2010. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. J. Neurosci. Methods 187 (2), 254–262.
Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143 (1), 29–36.
Harrell, F.J., 2001. Regression Modeling Strategies With Application to Linear Models, Logistic Regression, and Survival Analysis. Springer-Verlag, New York.
Harrell, F.J., 2013. rms: regression modeling strategies. R package version. 3 6–3 http://CRAN.R-project.org/package=rms.
Horan, W.P., Braff, D.L., Nuechterlein, K.H., Sugar, C.A., Cadenhead, K.S., Calkins, M.E., Dobie, D.J., Freedman, R., Greenwood, T.A., Gur, R.E., Gur, R.C., Light, G.A., Mintz, J., Olincy, A., Radant, A.D., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Green, M.F., 2008. Verbal working memory impairments in individuals with schizophrenia and their first-degree relatives: findings from the Consortium on the Genetics of Schizophrenia. Schizophr. Res. 103 (1–3), 218–228.
Iacono, W.G., 1998. Identifying psychophysiological risk for psychopathology: examples from substance abuse and schizophrenia research. Psychophysiology 35 (6), 621–637.
Jastak, S., Wilkinson, G., 1993. Wide Range Achievement Test-Revised 3. Jastak Associates, Wilmington, DE.
Johannesen, J.K., O'Donnell, B.F., Shekhar, A., McGrew, J.H., Hetrick, W.P., 2013. Diagnostic specificity of neurophysiological endophenotypes in schizophrenia and bipolar disorder. Schizophr. Bull. 39 (6), 1219–1229.
Johnson, R., Wichern, D., 2007. Applied Multivariate Statistical Analysis. sixth ed. Pearson Printice Hall, Upper Saddle River, NJ.
Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer-Verlag, New York, NY.
Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2 (3), 18–22.
Light, G., Greenwood, T.A., Swerdlow, N.R., Calkins, M.E., Freedman, R., Green, M.F., Gur, R.E., Gur, R.C., Lazzeroni, L.C., Nuechterlein, K.H., Olincy, A., Radant, A.D., Seidman, L.J., Siever, L.J., Silverman, J.M., Sprock, J., Stone, W.S., Sugar, C.A., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Braff, D.L., 2014. Comparison of the heritability of schizophrenia and endophenotypes in the COGS-1 family study. Schizophr. Bull. 40 (6), 1404–1411.
Maxwell, M., 1992. Family Interview for Genetic Studies (FIGS): a Manual for FIGS. Clinical Neurogenetics Branch, Bethesda, MD, NIMH.
Millard, S., 2013. EnvStats: an R Package for Environmental Statistics. Springer, New York.
Newcombe, R.G., 1998. Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat. Med. 17 (8), 873–890.
Nicodemus, K.K., Malley, J.D., Strobl, C., Ziegler, A., 2010. The behaviour of random forest permutation-based variable importance measures under predictor correlation. BMC Bioinf. 11, 110.
Nurnberger Jr., J.I., Blehar, M.C., Kaufmann, C.A., York-Cooler, C., Simpson, S.G., Harkavy-Friedman, J., Severe, J.B., Malaspina, D., Reich, T., 1994. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. Arch. Gen. Psychiatry 51 (11), 849–859 discussion 863-844.
Olincy, A., Braff, D.L., Adler, L.E., Cadenhead, K.S., Calkins, M.E., Dobie, D.J., Green, M.F., Greenwood, T.A., Gur, R.E., Gur, R.C., Light, G.A., Mintz, J., Nuechterlein, K.H., Radant, A.D., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Wagner, B.D., Freedman, R., 2010. Inhibition of the P50 cerebral evoked response to repeated auditory stimuli: results from the Consortium on Genetics of Schizophrenia. Schizophr. Res. 119 (1–3), 175–182.
Pepe, M., 2003. The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford University Press, Oxford, UK.
Pepe, M.S., Janes, H., Longton, G., Leisenring, W., Newcomb, P., 2004. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. Am. J. Epidemiol. 159 (9), 882–890.
Peters, R.M., Gjini, K., Templin, T.N., Boutros, N.N., 2014. A statistical methodology to improve accuracy in differentiating schizophrenia patients from healthy controls. Psychiatry Res. 216 (3), 333–339.
Price, G.W., Michie, P.T., Johnston, J., Innes-Brown, H., Kent, A., Clissa, P., Jablensky, A.V., 2006. A multivariate electrophysiological endophenotype, from a unitary cohort, shows greater research utility than any single feature in the Western Australian family study of schizophrenia. Biol. Psychiatry 60 (1), 1–10.
Radant, A.D., Dobie, D.J., Calkins, M.E., Olincy, A., Braff, D.L., Cadenhead, K.S., Freedman, R., Green, M.F., Greenwood, T.A., Gur, R.E., Light, G.A., Meichle, S.P., Mintz, J., Nuechterlein, K.H., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, M.T., Turetsky, B.I., Tsuang, D.W., 2007. Successful multisite measurement of antisaccade performance deficits in schizophrenia. Schizophr. Res. 89 (1–3), 320–329.
Radant, A.D., Dobie, D.J., Calkins, M.E., Olincy, A., Braff, D.L., Cadenhead, K.S., Freedman, R., Green, M.F., Greenwood, T.A., Gur, R.E., Gur, R.C., Light, G.A., Meichle, S.P., Millard, S.P., Mintz, J., Nuechterlein, K.H., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, M.T., Turetsky, B.I., Tsuang, D.W., 2010. Antisaccade performance in schizophrenia patients, their first-degree biological relatives, and community comparison subjects: data from the COGS study. Psychophysiology 47 (5), 846–856.
Radant, A.D., Millard, S.P., Braff, D.L., Calkins, M.E., Dobie, D.J., Freedman, R., Green, M.F., Greenwood, T.A., Gur, R.E., Gur, R.C., Lazzeroni, L.C., Light, G.A., Meichle, S.P., Nuechterlein, K.H., Olincy, A., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Sugar, C.A., Tsuang, M.T., Turetsky, B.I., Tsuang, D.W., 2015. Robust differences in antisaccade performance exist between COGS schizophrenia cases and controls regardless of recruitment strategies. Schizophr. Res. 163 (1–3), 47–52.
Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., Muller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf. 12, 77.
Rubin, D., 1987. Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons, New York.
Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014n. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511 (7510), 421–427.
Seidman, L.J., Hellemann, G., Nuechterlein, K.H., Greenwood, T.A., Braff, D.L., Cadenhead, K.S., Calkins, M.E., Freedman, R., Gur, R.E., Gur, R.C., Lazzeroni, L.C., Light, G.A., Olincy, A., Radant, A.D., Siever, L.J., Silverman, J.M., Sprock, J., Stone, W.S., Sugar, C., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Green, M.F., 2015. Factor structure and heritability of endophenotypes in schizophrenia: findings from the Consortium on the Genetics of Schizophrenia (COGS-1). Schizophr. Res. 163 (1–3), 73–79.
Stone, W.S., Giuliano, A.J., Tsuang, M.T., Braff, D.L., Cadenhead, K.S., Calkins, M.E., Dobie, D.J., Faraone, S.V., Freedman, R., Green, M.F., Greenwood, T.A., Gur, R.E., Gur, R.C., Light, G.A., Mintz, J., Nuechterlein, K.H., Olincy, A., Radant, A.D., Roe, A.H., Schork, N.J., Siever, L.J., Silverman, J.M., Swerdlow, N.R., Thomas, A.R., Tsuang, D.W., Turetsky, B.I., Seidman, L.J., 2011. Group and site differences on the California Verbal Learning Test in persons with schizophrenia and their first-degree relatives: findings from the Consortium on the Genetics of Schizophrenia (COGS). Schizophr. Res. 128 (1–3), 102–110.
Sun, G.W., Shook, T.L., Kay, G.L., 1996. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. J. Clin. Epidemiol. 49 (8), 907–916.
Swerdlow, N.R., Sprock, J., Light, G.A., Cadenhead, K., Calkins, M.E., Dobie, D.J., Freedman, R., Green, M.F., Greenwood, T.A., Gur, R.E., Mintz, J., Olincy, A., Nuechterlein, K.H., Radant, A.D., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Braff, D.L., 2007. Multi-site studies of acoustic startle and prepulse inhibition in humans: initial experience and methodological considerations based on studies by the Consortium on the Genetics of Schizophrenia. Schizophr. Res. 92 (1–3), 237–251.
Swerdlow, N.R., Gur, R.E., Braff, D.L., 2015. Consortium on the Genetics of Schizophrenia (COGS) assessment of endophenotypes for schizophrenia: an introduction to this Special Issue of Schizophrenia Research. Schizophr. Res. 163 (1–3), 9–16.
Thibaut, F., 2006. Schizophrenia: an example of complex genetic disease. World J. Biol. Psychiatry 7 (4), 194–197.
Thibaut, F., Boutros, N.N., Jarema, M., Oranje, B., Hasan, A., Daskalakis, Z.J., Wichniak, A., Schmitt, A., Riederer, P., Falkai, P., Markers, W.T.F.o.B., 2015. Consensus paper of the WFSBP Task Force on Biological Markers: criteria for biomarkers and endophenotypes of schizophrenia part I: neurophysiology. World J. Biol. Psychiatry 16 (5), 280–290.

Turetsky, B.I., Calkins, M.E., Light, G.A., Olincy, A., Radant, A.D., Swerdlow, N.R., 2007. Neurophysiological endophenotypes of schizophrenia: the viability of selected candidate measures. Schizophr. Bull. 33 (1), 69–94.

Turetsky, B.I., Greenwood, T.A., Olincy, A., Radant, A.D., Braff, D.L., Cadenhead, K.S., Dobie, D.J., Freedman, R., Green, M.F., Gur, R.E., Gur, R.C., Light, G.A., Mintz, J., Nuechterlein, K.H., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Calkins, M.E., 2008. Abnormal auditory N100 amplitude: a heritable endophenotype in first-degree relatives of schizophrenia probands. Biol. Psychiatry 64 (12), 1051–1059.

Zweig, M.H., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin. Chem. 39 (4), 561–577.