Multivariate prediction of emerging psychosis in adolescents at high risk for schizophrenia

Jai Shah[1],[2], **Shaun M Eack**[3],[4], **Debra M Montrose**[3], **Neeraj Tandon**[1], **Jean M Miewald**[3], **Konsale M Prasad**[3], and **Matcheri S Keshavan**[1],[2],[3],*

[1]Beth Israel Deaconess Medical Center, Boston, MA

[2]Department of Psychiatry, Harvard Medical School, Boston, MA

[3]Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, PA

[4]School of Social Work, University of Pittsburgh, Pittsburgh, PA

Abstract

Background—Accurate prediction of psychosis development in high-risk populations is an important but thus far elusive goal. Of the many diverse etiologic and risk factors identified thus far, few have been combined into prospective risk ascertainment models. We tested the predictive power of familial, neurobiological, socioenvironmental, cognitive and clinical risk factors through an integrative biopsychosocial model for emerging psychosis in young relatives at familial risk for schizophrenia.

Methods—96 young first- and second- degree relatives of schizophrenia probands were followed for an average of 2.38 (SD = 0.98) years to examine their trajectory towards psychosis. Iterative structural equation modelling utilizing multiple etiologic and risk factors was employed to estimate their joint contribution to prediction of psychosis development.

Results—The rate of conversion to psychosis over the study period was 12.5%. In the final model, clinical measures of schizotypy were directly predictive of conversion, with early (familial, biological, socioenvironmental) and cognitive risk factors indirectly predictive of psychosis through increased baseline clinical symptomatology. Our model provided an excellent fit to the observed data, with sensitivity of 0.17, specificity of 0.99, positive predictive value of 0.67 and negative predictive value of 0.89.

Conclusions—Integrative modeling of multivariate data from familial, neurobiological, socioenvironmental, cognitive and clinical domains represents a powerful approach to prediction of psychosis development. The high specificity and low sensitivity found using a combination of

*Corresponding Author: Matcheri S Keshavan, MD., Address: 6th Floor, Massachusetts Mental Health Center, 75 Fenwood Road, Boston, MA 02115. Phone: 617 754 1256. Fax: 617 754 1250. mkeshava@bidmc.harvard.edu.

Contributions

Jai Shah, Shaun Eack and Matcheri Keshavan designed the study. Jai Shah, Neeraj Tandon, Jean Miewald and Debra Montrose extracted data from charts and files. Jai Shah and Shaun Eack undertook statistical analyses. Jai Shah managed literature searches and wrote the first draft of the manuscript. All authors have contributed to and have approved the final manuscript.

Conflict of Interest

MSK has received grant support from GSK and Sunovion.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

such variables suggests that their utility may be in confirmatory testing among already selected high-risk individuals, rather than for initial screening. These findings also highlight the importance of data from a broad array of etiologic and risk factors, even within a familial high-risk population. With further refinement and validation, such methods could form key components of early detection, intervention and prevention programs.

Keywords

Psychosis; schizophrenia; high-risk; familial; prediction

Introduction

Accurate prediction of psychosis development during the premorbid and prodromal periods has long been sought by researchers and clinicians studying psychosis (Meares, 1959; Sullivan, 1994). If achieved, prediction could suggest early detection and targeted intervention strategies and might lead to substantial decreases in morbidity and burden of illness, and improvements in quality of life (Wyatt, 1991; McGorry et al., 2002).

While clinical impressions of expert mental health practitioners are thus far insufficient to reliably predict which members of a clinical high-risk group will transition to psychosis (Nelson and Yung, 2010), studies of familial high-risk (FHR) (Johnstone et al., 2003; Keshavan et al., 2004) and clinical high-risk (CHR) (McGorry et al., 2003; Addington et al., 2007) groups have gathered a wide range of relevant information. A critical issue is the need for accurate and robust techniques for prospectively identifying individuals at highest risk for conversion. Prediction has been complicated, however, by the multifactorial etiopathology of these illnesses and the breadth of nonspecific psychopathology that precedes psychosis (Keshavan et al., 2008).

Thus far, FHR and CHR prediction studies have investigated the role of cognitive, clinical and neurobiologic markers alongside family history. Eack et al (Eack et al., 2008) found that total brain volume, baseline neurocognitive deficits, and baseline psychosis proneness prospectively predicted emerging psychopathology development (rather than psychosis specifically) in a FHR population, with little overlap among these domains. In a larger North American population at CHR, Cannon et al calculated that familial risk with recent functional impairment, unusual thought content or suspicion/paranoia, social impairment and substance abuse were all predictive of later psychosis – with increased predictive power achieved via combinations of these variables (Cannon et al., 2008). This analysis was replicated in an independent Australian cohort, with similar but nonidentical results (Thompson et al., 2011). And in a European study combining CHR and cognitive "basic symptom" criteria, Ruhrmann et al (Ruhrmann et al., 2010) report high positive predictive value (PPV) for a six-variable model and introduce a prognostic index for assigning individual risk.

Psychobiological models utilized in these studies have generally had higher specificity than sensitivity, illustrating the need for further refinement. In this light, recent decades have seen increasing attention directed to the interaction of socioenvironmental factors with biological and psychological ones in psychosis (Faraone et al., 2002; van Os et al., 2005; Jarvis, 2007; Morgan et al., 2008; Tandon et al., 2008; Shah et al., 2011). Despite small effect sizes and low specificity, the widespread exposure to socioenvironmental risk factors suggests their substantial population-attributable risk (McGrath et al., 2004; Krabbendam and van Os, 2005; McGrath, 2006; Kirkbride et al., 2010). An initial multivariate FHR prediction study developed a model of interactive genetic and environmental factors (Carter et al., 2002);

subsequent investigations have been sparse, with mixed results, and primarily in CHR populations (Mason et al., 2004; Dragt et al., 2011).

Schizophrenia and related psychoses are increasingly considered illnesses of multifactorial etiology, with interactions between genetic and environmental (or individual-level and ecological-level) factors the subject of much investigation. The primary aim of this study was to explore interactions between available familial, social, environmental, cognitive, clinical and neurobiological markers in a young population of individuals at familial risk for psychosis. Given that diverse risk factors at various developmental stages have relevance for subsequent development of psychosis, models that take into account relationships between factors may offer a powerful approach for optimizing risk ascertainment. We hypothesized that iterative structural equation modelling (SEM) integrating distal and proximal risk factors from diverse domains would provide increased predictive power for psychosis development than do individual factors alone.

Materials and Methods

Participants

96 individuals with either first- or second-degree relatives diagnosed with schizophrenia or schizoaffective disorder were assessed using the Structured Clinical Interview for DSM-IV (SCID) (First et al., 2002) or Schedule for Affective Disorders and Schizophrenia-Child Version (K-SADS) instruments (Ambrosini et al., 1989). Diagnoses of the index cases through which high-risk subjects were recruited included schizoaffective disorder (34), schizophrenia paranoid type (19), schizophrenia residual type (5), and schizophrenia undifferentiated (27). A definitive diagnosis was not possible for 11 index cases; reasons included difficulties with English, the index case being deceased, index cases unwilling or unable to consent for interview, or the subject being adopted.

With respect to study participation, individuals with mental retardation, lifetime evidence of a psychotic disorder, prior exposure to antipsychotic medications, or significant neurological or medical conditions were excluded from entry. Those reporting recent substance abuse were also excluded due to the potential psychomimetic properties of many such substances. Although the age range for study inclusion was 8–25, mean age of included participants at the time of consent was 15.9 years (SD = 3.46), with an average of 9.43 (SD = 3.26) years of education (Table 1). Exactly half (48) were female; 43 (44.7%) were Caucasian, 51 (53.1%) were African-American, and 2 (2.1%) were Asian. Relatives were offspring, niece/nephews, siblings, grandchildren or aunts/uncles of affected individuals.

Procedures

Participants were family members of individuals with schizophrenia or schizoaffective disorder and were recruited through advertisements or referral by treating inpatient and outpatient clinicians in Pittsburgh, PA (Keshavan et al., 2004). SCID assessment for a diagnosis of schizophrenia or schizoaffective disorder in the index cases was first performed. Eligible study participants (*i.e.* relatives of index cases) provided initial historical data during an interview and received neuroimaging, neuropsychological and clinical assessments as described. Subsequent visits were attempted annually to evaluate for development of any significant clinical psychopathology using the SCID/K-SADS; participants were followed for an average of 2.38 (SD = 0.98) years. The Institutional Review Board of the University of Pittsburgh Medical Center approved this research. Written informed consent was acquired from subjects and/or parents or guardians prior to their participation in the study.

Familial risk

As there is no universally accepted standard for quantifying genetic or familial loading in high-risk studies, we utilized two separate measures. In our primary analysis, the subject's degree of relatedness to the affected family member (first- versus second-degree) was used as a simple index of risk, independent of the number of affected relatives. A supplementary analysis using a modified simplex/multiplex method (Glatt et al., 2006) was also undertaken. This approach assigns a value based on a combination of total number of affected relatives and degree of relatedness.

Socioenvironmental and ecological factors

These factors were recorded based on self-report, with exceptions noted below where hospital or clinical records were obtained. Where possible, reports from parents or guardians were sought. Urbanicity was measured as a continuous variable by neighborhood population density in the US zip code where subjects reported being raised during childhood/ adolescence, using data from the 2000 US Census (available at www.census.gov and http:// mcdc2.missouri.edu/). In cases where individuals were raised in more than one location, the zip code in which they lived for the longest period of time was recorded. Since those meeting criteria for cannabis or substance misuse at baseline were excluded (due to the potential psychotomimetic properties of illicit substances), subsequent self-report of development of cannabis abuse/dependence over the course of the study was recorded as follows: 0 = no crossover to abuse/dependence, 1 = crossover to cannabis abuse/dependence. Perinatal and/or birth complications were documented using a modified version of the Pregnancy History Instrument (Buka et al., 2000); hospital records, where possible, were accessed to verify paternal or caregiver reports. Childhood adversity was documented utilizing categorical variables in three dimensions: parental separation or loss of a parent; significant time (6 months or greater) spent outside of the parental home, including institutionalization in residential housing or jail, homelessness, or being cared for by friends or extended family members; and involvement of child/youth services agencies. Paternal age at conception was recorded per report of the subject, or where possible their parent/guardian.

Neuropsychological and psychopathological measures

The Revised Wechsler Adult Intelligence Scale (Wechsler, 1981), Wisconsin Card Sorting Test (Heaton et al., 1993), Spatial Working Memory Test (Cogtest, 2002), and a category/ letter fluency task (Benton and Hamscher, 1978) were utilized to assess IQ, executive functioning, working memory, and verbal fluency respectively. Chapman and colleagues' Perceptual Aberration, Magical Ideation, and Social Anhedonia scales (Chapman et al., 1978; Eckblad et al., 1982; Eckblad and Chapman, 1983) formed the basis of clinical assessments of psychosis proneness. Initial and follow-up interviews utilized SCID and K-SADS data to assess for clinical psychopathology development meeting DSM-IV criteria. Significant clinical psychopathology was logged in the case of (a) subjects who carried no baseline diagnosis but met criteria for an Axis I disorder based on SCID/K-SADS at follow-up, or (b) subjects who developed a more severe Axis I disorder between baseline and last follow-up. An example of the latter would be an anxiety disorder that later developed into a psychotic-spectrum disorder.

Image acquisition and processing

Structural MRI using a 1.5-T Signa whole body scanner and head coil (GE Medical Systems, Milwaukee, WI) was used to collect brain morphology data. Whole brain volumes were acquired in 124×1.5mm-thick continuous slices in the coronal plane, with spoiled gradient recalled acquisition in steady state pulse sequence (TE = 5ms, TR = 25ms, acquisition matrix = 256×192, FOV = 24cm). Manual checks for motional and quality were

independently conducted by trained research associates, and images were then normalized to standard MNI space and segmented using SPM5 (Wellcome Department of Cognitive Neurology, Institute of Neurology, London). During postprocessing, images were corrected for inhomogeneity artefacts using an algorithm for bias correction as part of the segmentation procedure, followed by smoothening using a 12mm Gaussian kernel. Total brain volumetric measurements were extracted using the Wake Forest University PickAtlas toolbox for SPM5 (Maldjian et al., 2003) to provide image masks. Regional definitions were acquired from Tzourio-Mazoyer and colleagues (Tzourio-Mazoyer et al., 2002).

Data analysis

Iterative structural equation modeling (SEM) was used to test an integrated predictive model of psychosis (schizophrenia, schizophreniform disorder, schizoaffective disorder, psychosis NOS) development utilizing the aforementioned factors. Significant clinical psychopathology was logged in the case of (a) subjects who carried no baseline diagnosis but met criteria for an Axis I disorder based on SCID/K-SADS at follow-up, or (b) subjects who developed a more severe Axis I disorder between baseline and last follow-up. An example of the latter would be an anxiety disorder that later developed into a psychotic-spectrum disorder. Presence of a baseline diagnosis was adjusted for by coding this variable in a categorical fashion: 0 = no baseline diagnosis; 1 = baseline Axis I diagnosis.

A full description of the application of SEM methodology to prediction of psychopathology development can be found in Eack et al. (Eack et al., 2008). Importantly, model design and testing in SEM draws on theory and previous data/evidence to represent and assess relationships between and among variables: measurement components of an SEM define putative latent variables based on streams of available data, while structural components of an SEM test hypothesized relationships. SEM can be particularly useful in situations where a higher-order construct is best represented not by one measured variable, but instead by multiple imperfectly measured variables that account for various aspects of the hypothetical construct.

Given the binary outcome variable under investigation here (whether subjects did or did not develop psychotic illness), Weighted Least Squares Mean and Variance adjusted estimation were utilized (Muthen and Muthen, 2001). Missing data were handled using Expectation Maximization estimates (Dempster et al., 1977) in order to reduce bias in covariance matrices (Schafer and Graham, 2002).

Overall statistics comparing the goodness-of-fit between observed data with hypothesized models tested the accuracy of the final model (Kline, 2011). Each of the Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), and χ^2 statistic (adjusted for categorical data) were calculated. Classification analyses were conducted to determine model accuracy using standard calculations for sensitivity, specificity, and positive and negative predictive values and likelihood ratios. Model-based predictions and confidence intervals were based on 1000 bootstrapped sample replications.

Results

Of the 96 individuals followed, 51 (53.1%) were found to have no baseline diagnosis while 45 (46.9%) had non-psychotic baseline diagnoses: 15 with attentional disorders; 4 with conduct disorders; 16 with mood disorders; 8 with anxiety disorders; and 2 with other. Twenty nine (30%) experienced an initial development (17; 17.7%) or worsening (12; 12.5%) of psychopathology on subsequent evaluations. Final diagnoses of the 96 study participants included 34 with no psychopathology, 15 with attentional disorders, 4 with conduct disorders, 18 with mood disorders, 10 with anxiety disorders, and 3 with other.

Twelve subjects (12.5%) transitioned to psychosis over the course of the study with diagnoses of schizoaffective disorder (4), psychosis NOS (4), schizophrenia (undifferentiated, 2; paranoid type, 1), and schizophreniform disorder (1). The age at which subjects converted to psychosis ranged from 13–24.

Integrative structural equation modeling was used to predict subsequent psychosis development after controlling for age, gender and baseline diagnostic status. Bivariate relationships among the variables used to construct this model are presented in Table 2.

Results from the final structural equation model (Figure 1) indicated that greater psychosis proneness at baseline (clinical assessment based on Chapman's rating scales for Magical Ideation (range 0–21), Social Anhedonia (range 0–26) and Perceptual Aberration (range 0–20)) was directly predictive of subsequent conversion to psychosis after adjusting for age, gender and baseline diagnostic status. Early risk factors recorded at baseline (familial risk, obstetric and prenatal complications, development of cannabis abuse/dependence, child welfare services involvement and removal from the family home for > 6 months) as well as cognitive dysfunction (IQ, perseverative and non-perseverative errors measuring executive function, and verbal fluency) were indirectly predictive of later psychosis through the mediating clinical measure.

As indicated in Figure 1, there was significant covariation between early risk factors and cognitive impairment. Some variables from each of the three domains – neurobiological (total brain volume via structural MRI), cognitive (spatial working memory), and socioenvironmental (urbanicity, paternal age at birth, loss of parent) – did not load onto the putative latent factors, and therefore did not further enhance predictive power for psychosis development in this sample.

Estimates of final model fit indicated excellent representation of the observed data, outcomes and relationships among the variables. In the primary analysis, χ^2 (32, N=96) = 34.4, $p = 0.35$; CFI = 0.987; RMSEA = 0.028. Model fit statistics for the simplex/multiplex analysis are listed in the appendix.

Classification analyses showed that 88% of subjects were correctly classified by this model in 1000 bootstrapped sample replications (Table 3). Model specificity (0.99) was excellent although sensitivity was low (0.17), indicating that the current model is stronger in identifying individuals who will remain healthy than those who will go on to develop psychosis.

Positive (0.67) and negative (0.89) predictive values and positive (+14) and negative (−0.84) likelihood ratios, which better account for population prevalence rates, were calculated. Based on 12.5% prevalence, individuals had a 67% chance (post-test OR 2.0) of developing psychosis if they were identified as doing so by the multivariate model, whereas those identified as not developing psychosis had an 11% chance (post-test OR 0.1) of developing psychosis.

A supplemental analysis was undertaken using a modified version of the simplex/multiplex method pioneered by Glatt et al (Glatt et al., 2006). Formal results of this analysis are presented in the Appendix.

Discussion

Although schizophrenia and related psychoses are increasingly being seen as disorders with multifactorial etiology, few studies thus far have evaluated the predictive power of integrative models incorporating a range of early and late risk factors from neurobiological,

socioenvironmental, cognitive and clinical domains. Through SEM, this study examined the relevant contribution of such factors to subsequent development of psychosis in adolescent at-risk relatives. Our findings suggest that baseline clinical psychosis proneness is directly predictive of subsequent transition to psychosis, while baseline neurocognitive impairment and early exposure to known familial and socioenvironmental risk factors are indirectly predictive of subsequent conversion through the mediating clinical measure.

This model used baseline and some longitudinal data to correctly classify 89% of individuals who did versus did not develop psychosis (Table 3). Given the low base-rate of conversion in this FHR population, the model's low sensitivity (0.17) is unsurprising. Statistics taking prevalence rates into account suggest that subjects identified by the model have a 67% chance of subsequently developing psychosis (+LR 14). Relative to an earlier analysis using SEM for overall psychopathology development without the inclusion of socioenvironmental factors (Eack et al., 2008), this study demonstrates improved specificity but lower sensitivity, comparable efficiency, and greater PPV. Relative to an initial FHR multivariate analysis for schizophrenia versus all other outcomes, our model has significantly lower sensitivity but improved PPV (Carter et al., 2002). This suggests its limited utility as an initial screening test. Due to its relatively high specificity (0.99), the current model may instead have more utility for testing of individuals already identified as being at especially high risk, or as a confirmatory test to be utilized in conjunction with, but subsequent to, instruments with high sensitivity. Because individuals identified in our multivariate model were fairly likely to develop psychosis, the model could also be particularly effective in highlighting those for whom preventative intervention would be high-yield.

Results of multivariate classification analyses presented here can be compared with statistics from three CHR populations (Cannon et al., 2008; Ruhrmann et al., 2010; Thompson et al., 2011). This study's FHR population was significantly younger, frequently pre-morbid (without significant psychopathology or functional impairment) and non-help seeking, and identified due to familial risk independent of symptomatic or functional status (Table 4). In conjunction with an earlier FHR report (Carter et al., 2002), our results suggest that prospective prediction utilizing baseline clinical, cognitive and biological data complemented by early individual and socioenvironmental history is a promising line of investigation. The relevance of distal risk markers for risk ascertainment – even within a familial high-risk population – is also notable. Sensitivity of 0.17, at the lower end of the range reported in CHR groups, is consistent with the lower base-rate of conversion in this study and the earlier and less stratified mental state of subjects (Keshavan et al., 2011).

Factors in the present model are depicted as staged, with early risk factors and cognitive impairment presenting early but mediated by the direct effects of psychosis proneness on eventual psychosis. While an alternative model of cognitive impairment and early risk factors acting independently on psychosis (without the mediating effects of psychosis proneness) was also investigated, such models provide an incomplete account of relationships between risk factors (see Figure 1, hashed lines). The importance of psychosis proneness is reinforced by the strong predictive power of a clinical index drawing on Chapman scales in the same population (Tandon et al., 2012). Thus, psychosis proneness might be considered as not simply a preceding risk factor, but the beginning clinical manifestations of the emerging psychotic disorder itself, of which cognitive, neurobiological, and socioenvironmental factors are directly predictive.

Findings in our final SEM should be interpreted in light of modest sample size. The robust conversion rates and (in two cases less than 16 years old) early age of conversion to psychosis may be influenced by the fact that many subjects were recruited from families with index cases under inpatient or intensive outpatient care, representative of significant

genetic and environmental loading and/or few protective factors. Nonetheless, 12.5% conversion is consistent with transition rates in independent studies of adolescents at FHR over a similar 2.5 year time-window (Johnstone et al., 2005). More individuals may develop psychosis in later years, although likely with lower frequency as they exit the period of greatest risk.

Since baseline cannabis abuse/dependence could not be utilized, subsequent development of cannabis misuse constituted the closest available marker of similar risk. Other well-known risk factors were incorporated into the initial model yet were found to have no significant contribution to predictive power. Their absence from our final SEM is potentially explained the relatively crude measures of social/environmental risk that were captured in this study and their smaller effect size. Common population-level factors such as urban upbringing may have been overwhelmed by limited variance and/or the high risk conferred by familial liability. Alternately, this might be due to theory-based decisions, for example the choice to code neighbourhood population density as continuous rather than categorical or otherwise non-linear (van Os et al., 2001). It is curious that total brain volume, which we have previously reported as relevant for prediction of general psychopathology development in a multivariate SEM (Eack et al., 2008), was noncontributory to psychosis prediction here. One explanation for this discrepancy is that overall brain volume is a relatively crude measurement of structural differences, and is insensitive to finer signals in specific regions.

Future refinements could draw upon additional risk and protective factors or differentially theorized relationships among variables. It is also conceivable that the constellation of factors (or the relationships among factors) relevant to accurate psychosis prediction will change based on the combination of early exposures and individual/ecological risks borne and experienced by particular individuals and populations at various developmental stages. Models could therefore be adjusted to account for the potentially staggered, latent, synergistic and/or cascading effects of risk and protective factors acting at critical time-points.

Conclusion

In addition to being important population-level contributors to psychosis development, our results indicate that certain socioenvironmental factors can contribute to individual-level risk prediction as well. Integrative modelling of multivariate data from a broad array of domains thus represents a powerful approach to prospective prediction of psychosis development: it can inform both high-risk and population-based strategies, and could become a key component of early detection, intervention and prevention programs.

Acknowledgments

References

Addington J, Cadenhead KS, Cannon TD, Cornblatt B, McGlashan TH, Perkins DO, Seidman LJ, Tsuang M, Walker EF, Woods SW, et al. North American Prodrome Longitudinal Study: a collaborative multisite approach to prodromal schizophrenia research. Schizophr Bull. 2007; 33:665–672. [PubMed: 17255119]

Ambrosini PJ, Metz C, Prabucki K, Lee JC. Videotape reliability of the third revised edition of the K-SADS. J Am Acad Child Adolesc Psychiatry. 1989; 28:723–728. [PubMed: 2793800]

Benton, A.; Hamscher, K. Multilingual aphasia examination manual (revised). Iowa City, IA: University of Iowa; 1978.

Buka SL, Goldstein JM, Seidman LJ, Tsuang MT. Maternal recall of pregnancy history: accuracy and bias in schizophrenia research. Schizophr Bull. 2000; 26:335–350. [PubMed: 10885635]

Cannon TD, Cadenhead K, Cornblatt B, Woods SW, Addington J, Walker E, Seidman LJ, Perkins D, Tsuang M, McGlashan T, et al. Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. Arch Gen Psychiatry. 2008; 65:28–37. [PubMed: 18180426]

Carter JW, Schulsinger F, Parnas J, Cannon T, Mednick SA. A multivariate prediction model of schizophrenia. Schizophr Bull. 2002; 28:649–682. [PubMed: 12795497]

Chapman LJ, Chapman JP, Raulin ML. Body-image aberration in Schizophrenia. J Abnorm Psychol. 1978; 87:399–407. [PubMed: 681612]

Cogtest. Cogtest: Computerized Cognitive Battery for Clinical Trials. 2002. Available at http://www.cogtest.com

Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data using the EM algorithm. J R Stat Soc B. 1977; 39:1–38.

Dragt S, Nieman DH, Veltman D, Becker HE, van de Fliert R, de Haan L, Linszen DH. Environmental factors and social adjustment as predictors of a first psychosis in subjects at ultra high risk. Schizophr Res. 2011; 125:69–76. [PubMed: 20884179]

Eack SM, Prasad KMR, Montrose DM, Goradia DD, Dworakowski D, Miewald J, Keshavan MS. An integrated psychobiological predictive model of emergent psychopathology among young relatives at risk for schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry. 2008; 32:1873–1878. [PubMed: 18812206]

Eckblad M, Chapman LJ. Magical ideation as an indicator of schizotypy. J Consult Clin Psychol. 1983; 51:215–225. [PubMed: 6841765]

Eckblad M, Chapman LJ, Chapman JP, Mishlove M. The revised social anhedonia scale. 1982 Unpublished.

Faraone SV, Brown CH, Glatt SJ, Tsuang MT. Preventing schizophrenia and psychotic behaviour: definitions and methodological issues. Can J Psychiatry. 2002; 47:527–537. [PubMed: 12211880]

First, M.; Spitzer, R.; Gibbon, M.; Williams, J. Structured clinical interview for DSM-IV-TR Axis I disorders, research version, patient edition. New York: Biometric Research, New York State Psychiatric Institute; 2002.

Glatt SJ, Stone WS, Faraone SV, Seidman LJ, Tsuang MT. Psychopathology, personality traits and social development of young first-degree relatives of patients with schizophrenia. Br J Psychiatry. 2006; 189:337–345. [PubMed: 17012657]

Heaton, R.; Chelune, G.; Talley, J.; Kay, G.; Curtis, G. Wisconsin Card Sorting Test Manual: revised and expanded. Odessa, FL: Psychological Assessment Resources Inc; 1993.

Jarvis GE. The social causes of psychosis in North American psychiatry: a review of a disappearing literature. Can J Psychiatry. 2007; 52:287–294. [PubMed: 17542379]

Johnstone EC, Ebmeier KP, Miller P, Owens DGC, Lawrie SM. Predicting schizophrenia: findings from the Edinburgh High-Risk Study. Br J Psychiatry. 2005; 186:18–25. [PubMed: 15630119]

Johnstone EC, Russell KD, Harrison LK, Lawrie SM. The Edinburgh High Risk Study: current status and future prospects. World Psychiatry. 2003; 2:45–49. [PubMed: 16946890]

Keshavan M, Montrose DM, Rajarethinam R, Diwadkar V, Prasad K, Sweeney JA. Psychopathology among offspring of parents with schizophrenia: relationship to premorbid impairments. Schizophr Res. 2008; 103:114–120. [PubMed: 18442896]

Keshavan MS, DeLisi LE, Seidman LJ. Early and broadly defined psychosis risk mental states. Schizophr Res. 2011; 126:1–10. [PubMed: 21123033]

Keshavan MS, Diwadkar VA, Montrose DM, Stanley JA, Pettegrew JW. Premorbid characterization in schizophrenia: the Pittsburgh High Risk Study. World Psychiatry. 2004; 3:163–168. [PubMed: 16633488]

Kirkbride J, Coid JW, Morgan C, Fearon P, Dazzan P, Yang M, Lloyd T, Harrison GL, Murray RM, Jones PB. Translating the epidemiology of psychosis into public mental health: evidence, challenges and future prospects. J Public Ment Health. 2010; 9:4–14. [PubMed: 21160544]

Kline, RB. Principles and Practice of Structure Equation Modeling. 3. New York: The Guilford Press; 2011.

Krabbendam L, van Os J. Schizophrenia and urbanicity: a major environmental influence--conditional on genetic risk. Schizophr Bull. 2005; 31:795–799. [PubMed: 16150958]

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 2003; 19:1233–1239. [PubMed: 12880848]

Mason O, Startup M, Halpin S, Schall U, Conrad A, Carr V. Risk factors for transition to first episode psychosis among individuals with "at-risk mental states". Schizophr Res. 2004; 71:227–237. [PubMed: 15474894]

McGorry PD, Yung AR, Phillips LJ. The "close-in" or ultra high-risk model: a safe and effective strategy for research and clinical intervention in prepsychotic mental disorder. Schizophr Bull. 2003; 29:771–790. [PubMed: 14989414]

McGorry PD, Yung AR, Phillips LJ, Yuen HP, Francey S, Cosgrave EM, Germano D, Bravin J, McDonald T, Blair A, et al. Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. Archives of General Psychiatry. 2002; 59:921–928. [PubMed: 12365879]

McGrath J. Variations in the incidence of schizophrenia: data versus dogma. Schizophr Bull. 2006; 32:195–197. [PubMed: 16135560]

McGrath J, Saha S, Welham J, Saadi ElO, MacCauley C, Chant D. A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. BMC Med. 2004; 2:13. [PubMed: 15115547]

Meares A. The diagnosis of prepsychotic schizophrenia. Lancet. 1959; 1:55–58. [PubMed: 13621637]

Morgan, C.; McKenzie, K.; Fearon, P. Society and Psychosis. Cambridge: Cambridge University Press; 2008.

Muthen, L.; Muthen, B. Mplus User's Guide. 2001.

Nelson B, Yung AR. Can clinicians predict psychosis in an ultra high risk group? Aust N Z J Psychiatry. 2010; 44:625–630. [PubMed: 20397780]

Ruhrmann S, Schultze-Lutter F, Salokangas RK, Heinimaa M, Linszen D, Dingemans P, Birchwood M, Patterson P, Juckel G, Heinz A, et al. Prediction of psychosis in adolescents and young adults at high risk: results from the prospective European prediction of psychosis study. Arch Gen Psychiatry. 2010; 67:241–251. [PubMed: 20194824]

Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods. 2002; 7:147–177. [PubMed: 12090408]

Shah J, Mizrahi R, McKenzie K. The four dimensions: a model for the social aetiology of psychosis. Br J Psychiatry. 2011; 199:11–14. [PubMed: 21719876]

Sullivan HS. The onset of schizophrenia. 1927. American Journal of Psychiatry. 1994; 151:134–139. [PubMed: 8192188]

Tandon N, Montrose D, Shah J, Rajarethinam RP, Diwadkar VA, Keshavan MS. Early prodromal symptoms can predict future psychosis in familial high-risk youth. J Psychiatr Res. 2012; 46:105–110. [PubMed: 22056319]

Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, "just the facts" what we know in 2008. 2. Epidemiology and etiology. Schizophr Res. 2008; 102:1–18. [PubMed: 18514488]

Thompson A, Nelson B, Yung A. Predictive validity of clinical variables in the "at risk" for psychosis population: international comparison with results from the North American Prodrome Longitudinal Study. Schizophr Res. 2011; 126:51–57. [PubMed: 21035313]

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 2002; 15:273–289. [PubMed: 11771995]

van Os J, Hanssen M, Bijl RV, Vollebergh W. Prevalence of psychotic disorder and community level of psychotic symptoms: an urban-rural comparison. Archives of General Psychiatry. 2001; 58:663–668. [PubMed: 11448373]

van Os J, Krabbendam L, Myin-Germeys I, Delespaul P. The schizophrenia envirome. Curr Opin Psychiatry. 2005; 18:141–145. [PubMed: 16639166]

Wechsler, D. Wechsler adult intelligence scale -- revised. New York: Psychological Corp; 1981.

Wyatt RJ. Neuroleptics and the natural course of schizophrenia. Schizophr Bull. 1991; 17:325–351. [PubMed: 1679255]

Appendix

In a supplemental analysis designed to measure an alternative index of familial risk, we have adopted a simplex/multiplex method previously utilized by Glatt et al. (Glatt et al., 2006) and modified slightly to utilize data available in the Pittsburgh sample. This method takes into account both degree of relatedness and the total number of affected relatives. In this scale, 0 = one affected second-degree relative, 1= one affected first degree relative, 2 = one affected first-degree relative *and* one affected second-degree relative, and 3 = at least two affected first-degree relatives. Results are presented in an intercorrelation table (Table 5) with a corresponding path analytic figure (Figure 2).

Individual prediction as to development/non-development of psychosis were identical to those of the primary (first-degree vs second-degree affected relative) analysis presented in the body of this article. Sensitivity, specificity, PPV and NPV therefore match those in Table 3.

In this secondary analysis as well, estimates of final model fit indicated excellent representation of the observed data, outcomes and relationships among the variables: χ^2 (31, N=96) = 34.8, p = 0.29; CFI = 0.978; RMSEA = 0.036.

Figure 1.
Final structural equation model predicting the emergence of psychosis in at-risk individuals based on clinical, cognitive, neurobiological and early risk data: χ^2 (32, N=96) = 34.4, p = 0.35; CFI = 0.987; RMSEA = 0.028. Path parameters are presented as standardized regression coefficients; $* p$ 0.05. Model estimates are adjusted for age, sex and baseline diagnostic status. Hashed lines indicate an alternate and incomplete model of cognitive impairment and early risk factors acting independently on psychosis development (without the mediating effect of psychosis proneness).
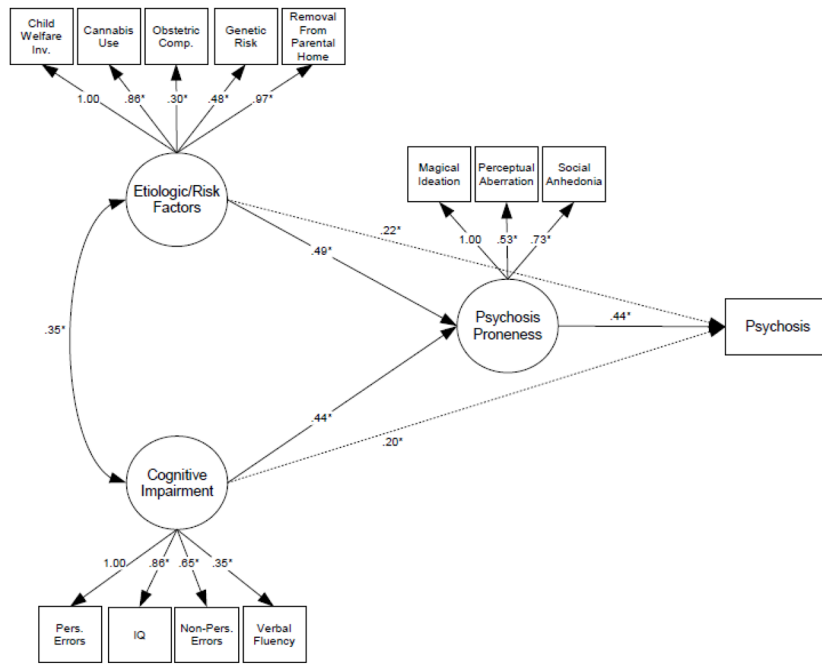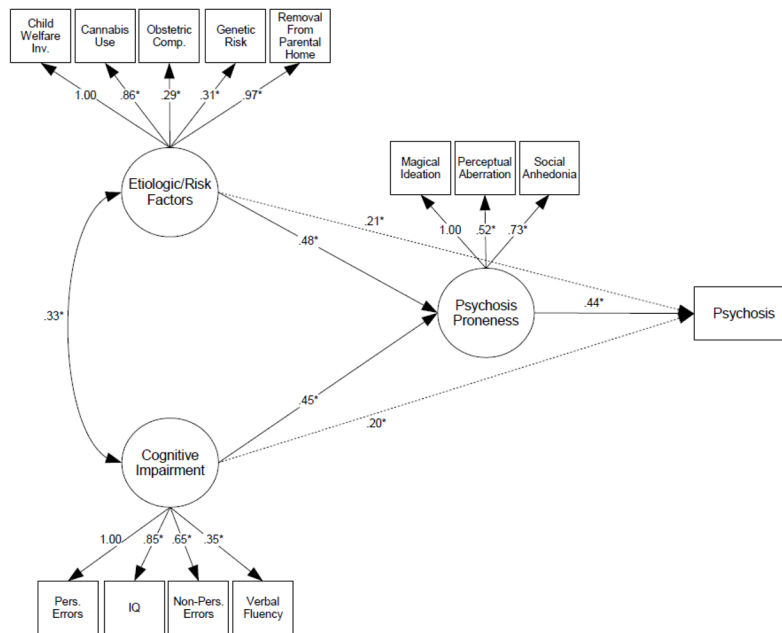
Figure 2.
Final structural equation model predicting the emergence of psychosis in at-risk individuals based on clinical, cognitive, neurobiological and early risk data: χ^2 (31, N=96) = 34.8, p = 0.29; CFI = 0.978; RMSEA = 0.036. Path parameters are presented as standardized regression coefficients; * p 0.05. Model estimates are adjusted for age, sex and baseline diagnostic status. Hashed lines indicate an alternate and incomplete model of cognitive impairment and early risk factors acting independently on psychosis development (without the mediating effect of psychosis proneness).

Table 1

Clinical and demographic characteristics of the familial high-risk sample.

	Converters (n = 12)	Nonconverters (n = 84)
Gender		
Male (n = 48)	5	43
Female (n = 48)	7	41
Age (mean, SD)	16.52, 3.17	15.84, 3.51
First-degree relatives		
Offspring	7	51
Siblings	2	14
Second-degree relatives	3	19
Race		
Caucasian	4	39
African-American	8	43
Latino/Hispanic	0	0
Asian	0	2

Table 2

Descriptive statistics and bivariate intercorrelations among neurocognitive, early/socioenvironmental and clinical study variables (N = 96)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
Cognition													
1. WCST: perseverative errors	-												
2. WCST: non-perseverative errors	**0.796**	-											
3. IQ	**0.438**	**0.522**	-										
4. Verbal fluency	**0.228**	**0.265**	**0.229**	-									
Early risk factors													
5. Child and youth services involvement	0.034	0.174	0.370	-0.142	-								
6. Removal from childhood home	-0.080	0.152	0.445	-0.042	**0.893**	-							
7. Development of cannabis abuse/dependence	-0.169	0.108	0.189	0.020	**0.579**	**0.852**	-						
8. Prenatal and birth complications	-0.023	0.012	0.099	-0.012	**0.422**	**0.151**	**0.359**	-					
9. Familial risk (0 = second-degree; 1 = first-degree)	0.297	0.357	0.279	0.162	**0.543**	**0.531**	**0.383**	**0.371**	-				
Clinical													
10. Chapman: Social anhedonia	0.205	0.317	0.354	0.266	0.417	0.471	0.477	0.077	0.064	-			
11. Chapman: Magical ideation	0.205	0.308	0.299	0.189	0.303	0.409	0.334	0.067	0.054	**0.432**	-		
12. Chapman: Perceptual aberration	0.132	0.223	0.255	0.202	0.290	0.365	0.181	-0.046	0.087	**0.319**	**0.691**	-	
13. Development of psychosis (0 = No; 1 = Yes)	-0.033	0.041	0.146	0.017	0.307	0.539	0.500	-0.229	-0.064	**0.300**	**0.375**	**0.267**	-
Mean	-0.12	-0.11	0.01	0.00	0.24	0.19	0.11	0.14	0.77	7.19	5.14	3.00	0.12
Standard deviation	1.14	1.10	1.01	1.00	0.43	0.39	0.32	0.13	0.42	4.97	3.89	3.87	0.33

Note: Intercorrelations are adjusted for age, sex and presence of baseline diagnosis. Within-domain relations appear in boldface.

Note: Cognitive variables have been scaled to z-scores, with mean of 0 and standard deviation 1.

Table 3

Prediction of psychosis development based on an integrative multivariate biopsychosocial structural equation model

Actual	Model-based prediction	
	Predicted to develop psychosis (3/96)	**Predicted to not develop psychosis (93/96)**
Developed psychosis (12/96)	2	10
Did not develop psychosis (84/96)	1	83

Note: Efficiency (95% CI intervals) = 0.88 (0.85 – 0.92); Sensitivity = 0.17 (−0.04 – 0.37); Specificity = 0.99 (0.96 – 1.01); PPV = 0.67 (0.06 – 1.25); NPV = 0.89 (0.87 – 0.92); +LR = 14; −LR = 0.84.

Confidence intervals are based on 1000 bootstrapped sample replications.

Table 4

Comparison of recent studies involving multivariate prediction of psychopathology or psychosis among clinical and familial high-risk groups

Reference	Cannon et al., 2008	Ruhrmann et al., 2010	Thompson et al., 2011	Dragt et al., 2011	Shah et al. (this study)
Location	Multi-site longitudinal North American study	Multi-site longitudinal European study	Longitudinal study in Melbourne, Australia	Longitudinal study in Amsterdam Netherlands	Longitudinal study in Pittsburgh, USA
Study Population	CHR	CHR/Basic Symptoms	CHR	CHR/Basic Symptoms	FHR
Sample Size	291	245	104	72	96
Age (SD)	18.1 (4.6)	23.0 (5.2)	19.4 (3.5)	19.3 (4.0)	15.9 (3.5)
Follow-up	Every 6 months to a maximum of 2.5 years	Every 9 months to a maximum of 1.5 years	Frequency unclear; maximum of 2.33 years	At 9, 18, 24 and 36 months	Annually to a maximum of 3 years
Outcome	Psychosis	Psychosis	Psychosis	Psychosis	Psychosis
Statistical methods	Cox proportional hazards, multivariate Cox regression	Cox proportional hazards, multivariate Cox regression	Cox proportional hazards, multivariate Cox regression	Cox proportional hazards, multivariate Cox regression	Iterative structural equation modelling
Risk of outcome (%)	35	19	34.6	26.4	12.5
Relevant Factors	Baseline combinations of: genetic risk with recent deterioration in functioning, levels of unusual thought content, suspicion/paranoia, social impairment, history of substance abuse	Baseline combinations of: positive symptoms, bizarre thinking, sleep disturbances, schizotypal disorder, level of functioning, years of education	Baseline combinations of: levels of unusual thought content, low functioning, genetic risk with recent deterioration in functioning	Baseline combinations of: urbanicity, social-sexual aspects, and personal adjustment	Baseline clinical measures of schizotypy (directly predictive); baseline early (familial, biological, socioenvironmental) and cognitive factors (indirectly predictive)
Sensitivity/Specificity	.10–.67/.84–.97	.417–.676/.456–.979	.174–.373/.872–.941	0.500–1.00/.694–.918	0.17/0.99
PPV/NPV	.46–.81/NR	.238–.833/.848–.870	.632–.667/.631–.682	NR/NR	0.67/0.89

CHR = clinical high-risk; FHR = familial high-risk; NR = not reported.

Table 5

Descriptive statistics and bivariate intercorrelations among neurocognitive, early/socioenvironmental and clinical study variables using simplex/multiplex method (N = 96)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
Cognition													
1. WCST: perseverative errors	-												
2. WCST: non-perseverative errors	**0.796**	-											
3. IQ	**0.438**	**0.522**	-										
4. Verbal fluency	**0.228**	**0.265**	**0.229**	-									
Early risk factors													
5. Child and youth services involvement	0.034	0.174	0.370	-0.142	-								
6. Removal from childhood home	-0.080	0.152	0.445	-0.042	**0.893**	-							
7. Development of cannabis abuse/dependence	-0.169	0.108	0.189	0.020	**0.579**	**0.852**	-						
8. Prenatal and birth complications	-0.023	0.012	0.099	-0.012	**0.422**	0.151	0.359	-					
9. Familial risk (Multiplex method)	0.159	0.218	0.130	0.080	**0.513**	0.341	0.302	0.093	-				
Clinical													
10. Chapman: Social anhedonia	0.205	0.317	0.354	0.266	0.417	0.471	0.477	0.077	-0.014	-			
11. Chapman: Magical ideation	0.205	0.308	0.299	0.189	0.303	0.409	0.334	0.067	-0.023	**0.432**	-		
12. Chapman: Perceptual aberration	0.132	0.223	0.255	0.202	0.290	0.365	0.181	- 0.046	0.057	**0.319**	**0.691**	-	
13. Development of psychosis (0 = No; 1 = Yes)	-0.033	0.041	0.146	0.017	0.307	0.539	0.500	-0.229	-0.167	**0.300**	**0.375**	**0.267**	-
Mean	-0.12	-0.11	0.01	0.00	0.24	0.19	0.11	0.14	1.24	7.19	5.14	3.00	0.12
Standard deviation	1.14	1.10	1.01	1.00	0.43	0.39	0.32	0.13	0.97	4.97	3.89	3.87	0.33

Note: Intercorrelations are adjusted for age, sex and presence of baseline diagnosis. Within-domain relations appear in boldface.

Note: Cognitive variables have been scaled to z-scores, with mean of 0 and standard deviation 1.