The “polyenviromic risk score”: Aggregating environmental risk factors predicts conversion to psychosis in familial high-risk subjects

Jaya L. Padmanabhan a,b, Jai L. Shah c,d, Neeraj Tandon a,e, Matcheri S. Keshavana,f,*

a Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA, USA
b Department of Behavioral Neurology and Neuropsychiatry, McLean Hospital, Belmont, MA, USA
c PEPP-Montréal, Douglas Mental Health University Institute, Montreal, Canada
d Department of Psychiatry, McGill University, Montreal, Canada
e Department of Psychiatry, Baylor College of Medicine, Texas Medical Center, Houston, TX, USA
f Department of Psychiatry, Harvard Medical School, Boston, MA, USA

Keywords:
Psychosis
Familial high risk
Schizophrenia
Prediction
Environment

ABSTRACT

Background: Young relatives of individuals with schizophrenia (i.e. youth at familial high-risk, FHR) are at increased risk of developing psychotic disorders, and show higher rates of psychiatric symptoms, cognitive and neurobiological abnormalities than non-relatives. It is not known whether overall exposure to environmental risk factors increases risk of conversion to psychosis in FHR subjects.

Methods: Subjects consisted of a pilot longitudinal sample of 83 young FHR subjects. As a proof of principle, we examined whether an aggregate score of exposure to environmental risk factors, which we term a ‘polyenviromic risk score’ (PERS), could predict conversion to psychosis. The PERS combines known environmental risk factors including cannabis use, urbanicity, season of birth, paternal age, obstetric and perinatal complications, and various types of childhood adversity, each weighted by its odds ratio for association with psychosis in the literature.

Results: A higher PERS was significantly associated with conversion to psychosis in young, familial high-risk subjects (OR = 1.97, p = 0.009). A model combining the PERS and clinical predictors had a sensitivity of 27% and specificity of 96%.

Conclusion: An aggregate index of environmental risk may help predict conversion to psychosis in FHR subjects.

*Corresponding author at: 75 Fenwood Ave., Boston, MA 02115, USA.
E-mail address: mkeshava@bidmc.harvard.edu (M.S. Keshavan).

1. Introduction

Several decades of epidemiological research have demonstrated the association of certain environmental variables with psychosis (van Os et al., 2010). These risk factors include urban birth or upbringing (Krabbendam and van Os, 2005), cannabis use (Kraan et al., 2016), season of birth (Davies et al., 2003), immigrant status (Bourque et al., 2011), paternal age (Torrey et al., 2009), obstetric or perinatal complications (Cannon et al., 2002; Geddes and Lawrie, 1995) and childhood adversity or abuse (Varese et al., 2012), among others.

In parallel with these advances in understanding environmental contributors to schizophrenia, identification of those at high risk of psychosis has become a priority. Several longitudinal studies of individuals at clinical high risk (CHR) or familial high risk (FHR) of schizophrenia have created psychosis prediction models, and a subset have evaluated whether environmental risk factors can enhance prediction of conversion to psychosis (Cannon et al., 2008; Ruhrmann et al., 2010; Shah et al., 2012; Tandon et al., 2012). Relatively few studies have evaluated the ability of environmental risk factors to predict psychosis risk among young subjects at FHR (i.e., first or second degree relatives of people with schizophrenia) (Johnstone et al., 2005; Shah et al., 2013), and these have reported mixed results. The Edinburgh High Risk Study, a longitudinal study of FHR individuals, found that substance abuse, but not obstetric complications or stressful life events, was associated with conversion to psychosis (Johnstone et al., 2005; McIntosh and Lawrie, 2001; Miller et al., 2006), while a study of FHR subjects in Denmark found that an unstable rearing environment predicted conversion to psychosis (Carter et al., 2002). Integrating a range of neuro-psychological, clinical, and environmental predictors into a structural equation model, our group has previously reported that cannabis use, obstetric complications, and removal from the parental home were indirectly predictive of conversion to psychosis (Shah et al., 2012). Other studies that assessed environmental risk factors did not retain them in final predictive models. For example, in a study of both CHR and FHR subjects, roughly 25% of whom had first or second degree relatives with psychosis (Ruhrmann et al., 2010), individual environmental risk factors did not significantly enhance prediction of conversion to psychosis and were therefore not included in the final predictive model.
However, it remains unknown whether an aggregate score representing loading for multiple environmental risk factors can predict conversion to psychosis in FHR individuals. Such an aggregate score may demonstrate whether cumulative exposure to environmental risk factors increases probability of conversion to psychosis, and could help identify FHR individuals at particularly elevated risk of psychosis. In this pilot longitudinal study of young FHR subjects, we calculated an index of overall loading for environmental risk factors for each subject, which we term the ‘polyenvironic risk score’ (PERS), analogous to the polygenic risk score used in genetics (Purcell et al., 2009). We evaluated the association of this score with conversion to psychosis over time and compared its predictive ability to a clinical model and to a combined clinical and PERS model. Finally, we present a sample web-based risk prediction application.

2. Methods

2.1. Subjects and clinical assessments

Subjects included 83 first or second degree relatives of people with schizophrenia or schizoaffective disorder, followed longitudinally for an average of 2.9 years. Exclusion criteria at the time of baseline evaluation included recent substance use, mental retardation, major neurological or medical conditions, and a history of psychosis or exposure to antipsychotic medication. Subjects were assessed with the Structured Clinical Assessment for DSM-IV (SCID) (First et al., 2002) or the Schedule for Affective Disorders and Schizophrenia-Child Version (K-SADS) (Ambrosini et al., 1989). This study was approved by the institutional review board of the University of Pittsburgh Medical Center, and was conducted in accordance with the Declaration of Helsinki. Baseline assessments included the Chapman schizotypy scales (Chapman et al., 1978; Eckblad and Chapman, 1983), the Scale of Prodromal Symptoms (Miller et al., 2003), the Wisconsin Card Sorting Test (Heaton et al., 1993), a go/no-go task, performance on the Identical Pairs (digits and shapes) version of the Continuous Performance Test (Cormblatt et al., 1988), a category/letter fluency task (Benton and Hamscher, 1978), and the Penn Emotion Recognition Task (Kohler et al., 2003).

2.2. Longitudinal follow-up

SCIDs were repeated annually in subjects until study completion, drop out from the study, or conversion to psychosis. Conversion to a psychotic disorder was determined during consensus conferences chaired by senior clinicians (MSK and others). These conferences involved review of medical charts and repeat SCID interviews, but did not incorporate research data, such as neuropsychological and clinical symptom scales. While environmental risk factors for psychosis were not explicitly reviewed during the consensus conferences, clinicians may have been exposed to this information either from medical charts or during the earlier research data collection. Repeat SCID interviews were not possible in all participants due to study attrition, which occurred due to subjects moving out of the area, being lost to follow-up or declining further participation in the study (Supplementary Table 1). However, all subjects who converted to psychosis did receive a confirmatory SCID, and there were no instances in which a SCID diagnosis of psychosis was rejected by the consensus conference.

2.3. Selection and binarization of environmental risk factors

An environmental risk factor was defined as an exposure (either physical, chemical, or infectious), a behavior pattern, or life event that could predispose an individual to schizophrenia (Ottman, 1996). We first reviewed our data and made a list of variables potentially meeting this definition of an environmental risk factor. To ensure that we were not missing other potential risk factors available in our data, we also conducted PubMed literature searches of English language meta-analyses and systematic reviews published in the last 5 years using the following search: (schizophrenia or psychosis) and (risk or epidemiology). We reviewed 8 meta-analyses or systematic reviews of environmental risk factors for schizophrenia (Akdeniz et al., 2014; Clarke et al., 2012; Davis et al., 2016; Hamlyn et al., 2013; Laurens et al., 2015; Matheson et al., 2011; Owen et al., 2016; Schmitt et al., 2014). We decided not to include sex or race because these factors may indirectly lead to different environmental exposures, but do not constitute environmental exposures in themselves. Immigrant status was not available in our data.

Second, for each possible risk factor on our list, we searched PubMed to identify meta-analyses, systematic reviews and original research reporting odds ratios for each risk factor’s association with psychosis, using the following search structure: (schizophrenia or psychosis) and (risk) and ([name of risk factor]). Manual review of articles was restricted to meta-analyses and systematic reviews when available. If no meta-analyses were available, systematic reviews and original research studies were evaluated. A risk factor was retained if meta-analyses or the majority of original research confirmed an association with psychosis (literature review is presented in Supplementary Table 2). We also examined the previously cited 8 reviews of environmental risk and schizophrenia for references.

Third, once it was decided to include a risk factor in the PERS, an odds ratio representing its association with psychosis or schizophrenia was selected from the literature. All odds ratios were obtained from meta-analyses, and more conservative odds ratios were usually selected when multiple options were available (details in Supplementary Table 2).

Thus, selection of environmental risk factors and odds ratios was based on several criteria: (1) the risk factor met the above definition of an environmental risk factor; (2) data for that risk factor was available in our sample; (3) substantial evidence in the literature supported the association of this risk factor with psychosis, defined as at least one meta-analysis or a consensus among systematic reviews and original research studies; and (4) an odds ratio could be obtained representing association with psychosis in a meta-analysis or original research.

Nine risk factors met these criteria. As will be described next, calculation of the PERS required binarization of each risk factor. Data collection and binarization of risk factors were conducted as follows:

1) Winter or spring birth: Date of birth was obtained at study enrollment. A date of birth between the winter and summer solstices was defined as a winter or spring birth, to match the definition used by the meta-analysis by Davies et al. (2003), which compared winter/spring to summer/fall birth.

2) Urbanicity: Zip codes of primary childhood place of residence were obtained from chart review. Population densities for these zip codes were obtained using year 2000 www.census.gov data. Urban locations were defined as having a population density of at least 1000 people per square mile per the US Census department definition of “urban” (https://www.census.gov/geo/reference/ua/urban-rural-2000.html). There were no meta-analytic odds ratios available by level of urbanicity, so we could not split this variable further by degree of urbanicity.

3) Cannabis use: As cannabis use was a baseline exclusion factor, subjects’ medical charts were reviewed for new onset of cannabis use during the course of the study. Because we did not collect data on severity of cannabis use, any level of cannabis use was binarized as being positive for cannabis abuse.

4) Advanced paternal age: This information was obtained by chart review or caregiver interview. Based on a meta-analysis by Torrey et al. (2009), paternal age was binarized at two levels: age ≥55 and <55, and age ≥55.

5) Obstetric and perinatal complications: Two sources of data on obstetric and perinatal complications were available: caregiver report via interview and the Pregnancy History Instrument (PHI) on a subset of individuals (Buka et al., 2000). Information from these two sources was
integrated into a single variable representing presence or absence of obstetric and perinatal complications. These included maternal diabetes, birth weight < 2500 g, emergency C-section, and bleeding in pregnancy, which are supported by meta-analysis (Cannon et al., 2002). However, we also included neonatal ICU stay, maternal infection, maternal substance abuse, and premature delivery (<4 weeks), among others. Due to the diversity of complications, the lack of meta-analytic odds ratios for certain types of complications, and the low frequency of any particular complication in this data set, we decided to use a global odds ratio from a meta-analysis (Geddes and Lawrie, 1995) for any obstetric/perinatal complication.

6–9): History of physical abuse, sexual abuse, neglect, and loss of a parent/parental separation were each coded as binary variables (present or absent) based on review of medical and research charts.

2.4. Statistical methods

Demographic and risk factors were tested for association with psychosis outcome using chi-squared or t-tests (Table 2).

The PERS was calculated by adapting a formula used for the calculation of polygenic risk scores (Purcell et al., 2009). To briefly summarize, the polygenic risk score uses an additive model to quantify an individual's genetic loading for a disorder, as conferred by multiple risk alleles. The odds ratios for the association of risk alleles with psychosis are derived from separate, large case-control genome-wide association studies. Individuals in the sample of interest are genotyped across single nucleotide polymorphisms. To calculate an individual's polygenic risk score, the log of the odds ratio for each risk allele is multiplied by the number of risk alleles an individual has at that location (either 0, 1 or 2). This is done for all single nucleotide polymorphisms that are included in the score. These products are added together and divided by the total number of single nucleotide polymorphisms included in the score.

In adapting the polygenic risk score formula to the PERS, we chose to binarize risk factors because meta-analyses typically present effect sizes as odds ratios for binarized variables or levels of a variable (as in parental age). Additionally, similar to the polygenic risk score, we decided to retain an additive model rather than incorporate interaction effects. While there is not conclusive evidence favoring an additive model, the evidence favoring an interactive model is also limited at this time, and we could not locate meta-analytic odds ratios for interaction effects.

The PERS for an individual was calculated as follows. Odds ratios for each environmental risk factor were obtained by literature review and meta-analytic odds ratios for interaction effects. We found that 10 of these factors had substantial supportive evidence in the literature from meta-analyses, of which one (childhood adversity) was a conglomerate of the others. The final PERS thus consisted of 9 risk factors (Table 1).

3. Results

Fourteen individuals developed a psychotic disorder by the end of the follow-up period, of which 6 had schizophrenia, 4 had schizoaffective disorder, 3 had psychosis NOS, and 1 had schizophreniform disorder.

Review of our available data yielded 14 possible variables meeting our definition of an environmental risk factor, of which 6 were various forms of childhood adversity or trauma (Supplementary Table 2). We found that 10 of these factors had substantial supportive evidence in the literature from meta-analyses, of which one (childhood adversity) was a conglomerate of the others. The final PERS thus consisted of 9 risk factors (Table 1).

Example: If an individual had a history of winter birth, urban upbringing, and obstetric complications, but no other risk factors, his PERS would be: (0.068 + 0.54 + 0.69) / 9 = 0.14.
Among clinical measures, the Chapman Magical Ideation scale (t = 2.2, p = 0.043), the Chapman Social Anhedonia scale (t = 2.1, p = 0.047), and performance on the Continuous Performance Task (digits) (t = 2.2, p = 0.041) were significantly correlated with psychosis (Supplementary Table 3), and were used to create a clinical predictive model (Table 3).

The PERS was significantly correlated with conversion to psychosis (OR = 1.97, p value = 0.009) and slightly enhanced prediction when added to a clinical model (Table 3). A sample risk prediction web application using a model containing only the PERS is available at https://kesh-lab.shinyapps.io/PERS-calc (note: this is provided strictly as an illustration and should not be used clinically).

Type of relationship (being a first versus second degree relative) was not a significant covariate in any model, and did not interact with PERS in predicting conversion to psychosis.

4. Discussion

In this pilot, proof-of-principle study, we created a measure of additive exposure to environmental risk factors, which we termed the polyenvironic risk score (PERS), and found that it was significantly correlated with conversion to psychosis in FHR subjects. Our findings support a ‘two hit’ or ‘multi-hit’ model of schizophrenia pathophysiology, whereby the combination of genetic and environmental risk can increase risk of psychosis (Bayer et al., 1999; Keshavan, 1999). Of note, this analysis focused on familial high risk individuals rather than those at clinical high risk of psychosis. Thus, the results of this predictive model cannot be generalized beyond the FHR population.

Despite the significant association with psychosis, the effect size for the PERS was about the same as reported odds ratios for individual risk factors from meta-analyses. This may have resulted from our small sample size or limitations in the accuracy of the PERS, as described below. Of note, odds ratios were mostly derived from studies conducted in the general population, not FHR individuals, and thus some environmental risk factors could have a lower effect size in the FHR population. It is also possible that aggregation of multiple risk factors does not significantly enhance prediction beyond individual risk factors, perhaps due to collinearity of risk factors. Meta-analyses of environmental risk factors rarely control for the presence of other environmental risk factors, so the effect sizes of individual environmental factors may be lower when concurrent risk factors are accounted for.

Relatively few studies have examined the additive effect of multiple environmental factors on risk of psychosis. In a sample of individuals with schizophrenia, Stepien et al. reported an association between increased number of environmental risk factors and age at schizophrenia onset, such that individuals who had been exposed to 4 or more risk factors had a significantly lower age of onset than those exposed to 3 factors, and so on (Stepniak et al., 2014). Cougnard et al. reported that additive exposure to three risk factors - cannabis use, childhood trauma, and urbanicity - interacted with baseline psychotic experiences in predicting persistent psychotic symptoms three years later in the general population (Cougnard et al., 2007). Although not examining conversion to a psychotic disorder, these studies imply a possible additive effect of multiple environmental risk factors on psychosis risk. Additionally, the aggregate index used in these studies was the total number of risk factors, rather than a weighted sum. However, to our knowledge, the association of cumulative environmental risk with conversion to psychosis in an FHR population has not been previously demonstrated, and no study has used an aggregate index weighted by the effect sizes of each risk factor.

There were several limitations to this study. First, and most importantly, although we kept the PERS as specific and precise as our available data and the literature would allow, the PERS construct was still limited...
in some respects. For example, cannabis and urbanicity are believed to have ‘dose dependent’ effects on psychosis risk (Marconi et al., 2016; Vassos et al., 2012), but we did not have data on severity of cannabis use in our subjects, and meta-analytic odds ratios for different levels of urbanicity were not available. Additionally, some known risk factors, such as immigrant status (Bourque et al., 2011), were not collected in this sample.

A second limitation is that the PERS assumes an additive relationship between environmental risk factors, and does not account for interaction effects. However, while there is not conclusive evidence favoring an additive model, evidence is also limited regarding an interactive model. Thus far, some research indicates an interaction between trauma and cannabis use on psychosis risk (Harley et al., 2010; Houston et al., 2008), but other studies have not replicated these findings (Kuepper et al., 2011; Sideli et al., 2015), and there are no meta-analytic odds ratios available by which to weight a presumed interaction effect in the PERS.

Third, while environmental risk factors were not reviewed during consensus conferences to determine conversion to psychosis, clinicians on the consensus committee may have been exposed to this information through medical record review or previous research assessments, which in theory could have biased their diagnostic determinations. However, we believe the risk of a significant bias is relatively low, because our current analysis was not a primary aim of the original study. Finally, this was a proof-of-principle study with a small sample size and is best viewed as an initial application of a method. Replication is needed in larger samples to determine accurate correspondence between scores and psychosis risk.

In addition to validating the PERS or similar constructs in larger samples, future studies could assess interactions of the PERS with neurobiological risk factors, such as changes in brain structure (Koutsouleris et al., 2009), and psychological phenomena, such as negative emotions (Garety et al., 2001; Thewissen et al., 2011). While neuropsychological and neurobiological measures have advanced our understanding of psychosis risk, environmental factors are critically important in the development of psychosis, especially at a population level. Use of a PERS or a similar construct may help disentangle the complex threads of psychosocial phenomena, such as negative emotions and psychological risk factors, such as changes in brain structure (Koutsouleris et al., 2011; Sideli et al., 2015), and there are no meta-analytic odds ratios available by which to weight a presumed interaction effect in the PERS.


