Polygenic Risk Score, Parental Socioeconomic Status, Family History of Psychiatric Disorders, and the Risk for Schizophrenia
A Danish Population-Based Study and Meta-analysis

Esben Agerbo, DrMedSc; Patrick F. Sullivan, FRANZCP; Bjarni J. Vilhjálmsson, PhD; Carsten B. Pedersen, DrMedSc; Ole Mors, PhD; Anders D. Berglum, PhD; David M. Hougaard, DrMedSc; Mads V. Hollegaard, PhD; Sandra Meier, PhD; Manuel Mattheisen, PhD; Stephan Ripke, MD; Naomi R. Wray, PhD; Preben B. Mortensen, DrMedSc

**IMPORTANCE** Schizophrenia has a complex etiology influenced both by genetic and nongenetic factors but disentangling these factors is difficult.

**OBJECTIVE** To estimate (1) how strongly the risk for schizophrenia relates to the mutual effect of the polygenic risk score, parental socioeconomic status, and family history of psychiatric disorders; (2) the fraction of cases that could be prevented if no one was exposed to these factors; (3) whether family background interacts with an individual's genetic liability so that specific subgroups are particularly risk prone; and (4) to what extent a proband's genetic makeup mediates the risk associated with familial background.

**DESIGN, SETTINGS, AND PARTICIPANTS** We conducted a nested case-control study based on Danish population-based registers. The study consisted of 866 patients diagnosed as having schizophrenia between January 1, 1994, and December 31, 2006, and 871 matched control individuals. Genome-wide data and family psychiatric and socioeconomic background information were obtained from neonatal biobanks and national registers. Results from a separate meta-analysis (34 600 cases and 45 968 control individuals) were applied to calculate polygenic risk scores.

**EXPOSURES** Polygenic risk scores, parental socioeconomic status, and family psychiatric history.

**MAIN OUTCOMES AND MEASURES** Odds ratios (ORs), attributable risks, liability $R^2$ values, and proportions mediated.

**RESULTS** Schizophrenia was associated with the polygenic risk score (OR, 8.01; 95% CI, 4.53-14.16 for highest vs lowest decile), socioeconomic status (OR, 8.10; 95% CI, 3.24-20.3 for 6 vs no exposures), and a history of schizophrenia/psychoses (OR, 4.18; 95% CI, 2.57-6.79). The $R^2$ values were 3.4% (95% CI, 2.1-4.6) for the polygenic risk score, 3.1% (95% CI, 1.9-4.3) for parental socioeconomic status, and 3.4% (95% CI, 2.1-4.6) for family history. Socioeconomic status and psychiatric history accounted for 45.8% (95% CI, 36.1-55.5) and 25.8% (95% CI, 21.2-30.5) of cases, respectively. There was an interaction between the polygenic risk score and family history ($P = .03$). A total of 17.4% (95% CI, 9.1-26.6) of the effect associated with family history of schizophrenia/psychoses was mediated through the polygenic risk score.

**CONCLUSIONS AND RELEVANCE** Schizophrenia was associated with the polygenic risk score, family psychiatric history, and socioeconomic status. Our study demonstrated that family history of schizophrenia/psychoses is partly mediated through the individual's genetic liability.
Schizophrenia is a debilitating and complex disorder with a lifetime risk of approximately 1%. Familial, adoption, twin, and sibling studies have consistently shown that schizophrenia is heritable, with a substantial overlap with other psychiatric disorders. Genetic evidence for specific variants have been emerging, as 108 schizophrenia-associated loci have been identified by the Psychiatric Genomics Consortium. Despite the small effect size of single loci, together the genome-wide-significant loci were estimated to explain 3.4% of the variance in liability, and the cumulative effect of common loci expressed as a polygenic risk score (PRS) was estimated to explain 7% of the variance in liability.

Environmental risk factors have long been recognized to play a role in the etiology of schizophrenia. These include various factors such as place and season of birth, maternal obstetrical complications, parental age, neonatal vitamin D levels, prenatal infection (e.g., influenza, toxoplasmosis, and herpes simplex virus), and low social class, where the latter appears to be both a cause and consequence of schizophrenia.

Gene-environment interactions may be important in the etiology of schizophrenia. Studies have used psychiatric family history as a proxy for genetic liability rather than actual genetic variation, as perhaps no data set exists where gene-environment interactions are identifiable. To our knowledge, no study has taken the gene-environment hypothesis further to estimate the proportion of an upstream factor that is mediated through the genetic liability for schizophrenia.

We used Denmark’s population-based registers, the Danish Neonatal Screening Biobank, and separate metadata from the largest published schizophrenia genome-wide association study to pursue the following questions: (1) how strongly is the risk for schizophrenia related to the mutual effect of the PRS, parental socioeconomic status, and family history of psychiatric disorders? (2) In theory, what fraction of cases could be prevented if no one was exposed to these factors? (3) Do familial backgrounds interact with an individual’s genetic liability so that specific subgroups of individuals are particularly risk prone? (4) How much of an excess risk associated with familial background is mediated through the offspring’s genetic makeup?

Methods

Data were obtained by linking Danish population-based registers using the unique personal identification number, which is assigned to all individuals who have been resident in Denmark since 1968 and used across all registration systems. The Danish Civil Registration System contains dates of birth, deaths, emigrations, and links to family members. The Danish Neonatal Screening Biobank stores dried blood spot samples collected at birth from nearly all infants born in Denmark after 1981. The Psychiatric Central Research Register includes all admission dates and International Classification of Diseases, Eighth Revision (ICD-8) and International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) diagnoses. It also covers all psychiatric inpatient facilities since 1969 and outpatient contacts since 1994. All diagnoses are based on clinical diagnoses assigned by physicians at discharge and the diagnosis of schizophrenia has been validated with good results. The Integrated Database for Longitudinal Labour Market Research covers the entire population and contains yearly information from 1980 including income, marital status, education, and birth place. The study was approved by the Danish Data Protection Agency; patient consent was waived as identities were blinded to the investigators.

We studied all singleton births since 1981 with a DNA sample available from the neonatal biobank and who had been given a diagnosis in the psychiatric register with an ICD-10 F20 code for schizophrenia between January 1, 1994, and December 31, 2006. Each case was matched with a randomly selected control individual of the same sex and with the same birthday. A control individual was only eligible provided he or she was born and resident in Denmark and not diagnosed as having schizophrenia before the date the case received a diagnosis. However, none of the control individuals were diagnosed as having schizophrenia within the study period. Parents and maternal siblings were identified using the Civil Registration System. DNA was extracted from the dried blood samples, whole-genome amplified (in triplicate using the Qia-gen REPLI-g mini kit and the 3 separate reactions were pooled), and genotyped with Illumina Human 610-Quad BeadChip array. Genotyping and quality-control details have been published previously.

We conducted a meta-analysis of association results of all Psychiatric Genomics Consortium samples after excluding the Danish study participants (discovery sample of 34 600 cases and 45 968 control individuals). We retained single-nucleotide polymorphisms (SNPs) that had minor allele frequency greater than 1% and imputation information score greater than 0.6 in both the discovery and Danish samples. Missing SNPs were imputed using the 1000 Genomes Project reference panel (release version 3 macGT1). We selected approximately 100 000 SNPs, removing SNPs with linkage disequilibrium with an R² value greater than 0.1 in 500-kb windows and preferentially retaining SNPs that were most associated in any region. We generated the PRS in the Danish sample based on P value cutoffs (the eAppendix in the Supplement contains details). A PRS is a sum of schizophrenia risk alleles carried by an individual, where each term is weighted by the corresponding log-odds ratio from the discovery sample. The weights are from a completely separate sample. In this study and a previous study by the Schizophrenia Working Group of the Psychiatric Genomics Consortium, P < .05 was used to achieve a balance between the number of false-positive and true-positive risk alleles (the eAppendix in the Supplement contains details for choosing this threshold). All analyses were repeated for thresholds of P < .01 and P < .10 (eAppendix in the Supplement).

We generated a parental socioeconomic status score for the year prior to the patient’s birth as a sum of 6 previously confirmed risk factors: father’s or mother’s gross income in the lowest quintile, father or mother being unemployed or otherwise outside the labor market, and father’s or mother’s highest educational level less than high school completion. Each
Variables Associated With Schizophrenia Risk

Original Investigation Research

Risk factor was scored as a binary (1 = yes and 0 = no) and the socioeconomic status was an equally weighted sum of the scores.

The psychiatric register was used to extract explanatory variables indicating whether the patient’s parents or siblings had been given a diagnosis according to the hierarchy: schizophrenia or related psychosis (ICD-8 codes 295, 297, 298.39, and 301.83, and ICD-10 codes F20-F29), bipolar affective disorder (ICD-8 codes 296.19, 296.39, 296.29, 296.89, 296.99, 296.09, 298.09, 298.19, 300.49, and 301.19, and ICD-10 codes F33-F34, F38, and F39), or any other psychiatric disorder before the matching date (details in the study by Mortensen et al\(^{18}\)). However, we are unaware of any register-based studies that have validated these diagnostic classifications.\(^{11}\)

Initial analyses were conducted with logistic regression. All analyses were adjusted by regression for sex, birth year (5 categories), and ancestry (using the first 10 genomic principal components), and adjusted odds ratios are also mutually adjusted.

Results

The sample comprised 866 cases with schizophrenia and 871 control individuals. Table 1 shows the ORs associated with the PRS, parental socioeconomic status, and family history of psychiatric disorders, as well as the number of exposed individuals. The risk for schizophrenia increased steadily with the PRS (ie, with increasing estimated liability to schizophrenia).

Figure 1 shows the ORs for schizophrenia vs the continuous PRS. The Akaike Information Criterion and the likelihood ratio test (P < .001) indicate that the continuum model is superior to the decimalized model. Moreover, Table 1 shows that the risk was increasing with the degree of parental socioeconomic disadvantage, with ORs ranging between 1.56 (95% CI, 1.18-2.03) and 8.10 (95% CI, 3.24-20.3) compared with individuals without any parental socioeconomic disadvantage.

Table 1. Odds Ratios for Schizophrenia in Relation to the Polygenic Risk Score, Parental Socioeconomic Status at Birth, and Family Psychiatric History

<table>
<thead>
<tr>
<th>Risk Factor and Category</th>
<th>No. of Cases/ Controls</th>
<th>Odds Ratio (95% CI)</th>
<th>Unadjusted</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polygenic risk score, decile(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest</td>
<td>130/43</td>
<td>8.01 (4.53-14.16)</td>
<td>7.36 (4.07-13.3)</td>
<td></td>
</tr>
<tr>
<td>Ninth</td>
<td>111/63</td>
<td>3.70 (2.36-5.78)</td>
<td>3.36 (2.11-5.35)</td>
<td></td>
</tr>
<tr>
<td>Eighth</td>
<td>92/82</td>
<td>2.31 (1.49-3.58)</td>
<td>2.15 (1.37-3.39)</td>
<td></td>
</tr>
<tr>
<td>Seventh</td>
<td>94/80</td>
<td>2.45 (1.58-3.79)</td>
<td>2.30 (1.46-3.62)</td>
<td></td>
</tr>
<tr>
<td>Sixth</td>
<td>83/91</td>
<td>1.88 (1.22-2.91)</td>
<td>1.67 (1.06-2.63)</td>
<td></td>
</tr>
<tr>
<td>Fifth</td>
<td>73/100</td>
<td>1.49 (0.96-2.32)</td>
<td>1.52 (0.96-2.40)</td>
<td></td>
</tr>
<tr>
<td>Fourth</td>
<td>74/100</td>
<td>1.50 (0.97-2.32)</td>
<td>1.50 (0.96-2.37)</td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td>78/96</td>
<td>1.66 (1.07-2.57)</td>
<td>1.42 (0.90-2.24)</td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>74/100</td>
<td>1.52 (0.98-2.36)</td>
<td>1.50 (0.95-2.36)</td>
<td></td>
</tr>
<tr>
<td>Lowest</td>
<td>57/116</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td></td>
</tr>
<tr>
<td>Parental socioeconomic status, No. of risk factors(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>26/6</td>
<td>8.10 (3.24-20.3)</td>
<td>5.58 (2.16-14.4)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>50/25</td>
<td>3.39 (1.99-5.77)</td>
<td>2.18 (1.35-4.20)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>96/64</td>
<td>2.59 (1.76-3.80)</td>
<td>2.05 (1.36-3.07)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>125/82</td>
<td>2.84 (2.01-4.03)</td>
<td>2.35 (1.63-3.38)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>217/166</td>
<td>2.45 (1.84-3.27)</td>
<td>2.14 (1.58-2.89)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>199/241</td>
<td>1.56 (1.18-2.05)</td>
<td>1.48 (1.11-1.97)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>153/287</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td></td>
</tr>
<tr>
<td>Family psychiatric history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia or other psychoses(^d)</td>
<td>78/23</td>
<td>4.18 (2.57-6.79)</td>
<td>2.60 (1.56-4.31)</td>
<td></td>
</tr>
<tr>
<td>Bipolar affective disorders</td>
<td>78/39</td>
<td>2.82 (1.88-4.22)</td>
<td>2.46 (1.62-3.73)</td>
<td></td>
</tr>
<tr>
<td>Other psychiatric disorders</td>
<td>184/96</td>
<td>2.60 (1.98-3.43)</td>
<td>2.31 (1.73-3.08)</td>
<td></td>
</tr>
<tr>
<td>No psychiatric disorder</td>
<td>526/713</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) The odds ratios are adjusted for sex, birth year (5 categories), and ancestry (using the first 10 genomic principal components), and adjusted odds ratios are also mutually adjusted.

\(^b\) The boundaries that define the 10 deciles are 413.0, 418.5, 419.7, 420.5, 421.2, 421.8, 422.5, 423.1, 424.1, 425.7, and 437.2.

\(^c\) Socioeconomic status is generated as a sum of 6 factors described in the Methods section.

\(^d\) Only 8 control individuals (52 cases) had a first-degree relative with a defined schizophrenia that results in an odds ratio of 8.27 (95% CI, 3.87-17.66).

Downloaded From: http://archpsyc.jamanetwork.com/ by a University of St. Andrews Library User on 05/18/2015
risk factors. The ORs related to the factor that constitutes parental socioeconomic status are shown in the eAppendix in the Supplement. As in prior studies, Table 1 shows that familial histories of psychiatric disorders were highly predictive of schizophrenia. The ORs were only slightly attenuated after being fully and mutually adjusted, with the exceptions being the ORs associated with a family history of schizophrenia/psychoses and being exposed to more than 6 parental socioeconomic risk factors.

The population-attributable risk is an estimate of the fraction of cases that would be prevented if all individuals had the risk of those not exposed. Table 2 shows the attributable risk associated with parental socioeconomic status (45.8%; 95% CI, 36.1-55.5) and familial history of psychiatric disorders (25.8%; 95% CI, 21.2-30.5). Thus, a sizeable proportion of cases can be attributed to these 2 factors. The attributable risk was not estimated for the PRS because no obvious reference exists. The liability $R^2$ value measures how much of the variation in liability to disease in a population is explained by the variation in a risk factor. As expected, the $R^2$ values were considerably smaller for parental socioeconomic status (3.1%; 95% CI, 1.9-4.3), family history (3.4%; 95% CI, 2.1-4.6), and the PRS (3.4%; 95% CI, 2.1-4.6), and their mutual effect was 7.8% (95% CI, 5.9-9.6).

Because genetic liability is inherited and because a broad class of psychiatric disorders confer an increased risk, we estimated the percentage of family history of psychiatric disorders mediated through the PRS (Table 3). Our analyses suggested that 17.4% (95% CI, 9.1-26.6) of the effect associated with a family history of schizophrenia/psychoses was mediated through the PRS. The interaction between the PRS and family history of schizophrenia/psychoses was significant ($P = .03$).

Taking this interaction into account implied that the proportion mediated increased to 47.7% (95% CI, 9.9-93.5). However, this proportion must be interpreted with caution because of the wide confidence interval. The direct and indirect ORs were 1.87 (95% CI, 0.68-3.28) and 1.77 (95% CI, 1.07-2.67), respectively (eAppendix in the Supplement).

After modeling the interactions, the main effect associated with familial history of schizophrenia/psychoses decreased to 2.71 (95% CI, 1.58-4.64) and the OR increased with 3.26 (95% CI, 1.74-6.10) per-PRS standard deviation, whereas the increase was 1.59 (95% CI, 1.38-1.83) among those with no history. To further elucidate the interaction, the ORs associated with the interaction between the continuous PRS and family history are depicted in Figure 2. The combination of no history and the lowest PRS quartile is the reference. Although
there were few exposed control individuals, Figure 2 suggests a diminished main effect of a family history of schizophrenia/psychoses at the lowest PRS level; however, the OR increases rapidly across the highest 3 quartiles of the PRS. Thus, family history of schizophrenia/psychoses has limited impact on the schizophrenia risk among individuals with low genetic liability but the impact increases with increasing liability. The ORs associated with family histories of schizophrenia/psychoses and bipolar affective disorders at the lowest level of the PRS were statistically indistinguishable ($P = .46$). Analogous figures for the other 9 PRSs are shown in the eAppendix in the Supplement.

Our data provided little evidence of parental socioeconomic status being mediated through the PRS (28.0%; 95% CI, −13.2 to 70.0; $P = .25$) or to suggest that the 2 factors interacted ($P = .28$).

Discussion

We found that the risk for schizophrenia in a population-based sample was strongly associated with an individual’s PRS, parental socioeconomic status, and family history of psychiatric disorders. Only modest fractions of the variation in liability were explained by the variation in the PRS (3.4%; 95% CI, 2.1-4.6), parental socioeconomic status (3.1%; 95% CI, 1.9-4.3), and family history (3.4%; 95% CI, 2.1-4.6). A sizeable percentage of cases was attributed to parental socioeconomic status (45.8%; 95% CI, 36.1-55.5) and familial history of psychiatric disorders (25.8%; 95% CI, 21.2-30.5). Our analyses showed that 17.4% (95% CI, 9.1-26.6) of the effect associated with a family history of schizophrenia/psychoses was mediated through the PRS.

Previous studies have found a positive correlation between the PRS and the risk for schizophrenia, recognizing that the PRS cannot be used as a diagnostic test for schizophrenia because only a modest fraction is explained by the variation in the PRS. Nevertheless, the PRS will probably become an important tool in uncovering the etiology of schizophrenia. For example, one study already suggests that the PRS was increased among treatment-resistant patients. The risk for schizophrenia was more closely linked to the PRS than to socioeconomic status and family history, although all 3 factors were highly indicative of schizophrenia.

Our present analyses suggest that family history of schizophrenia/psychoses has only limited impact on the risk for schizophrenia among individuals with low genetic liability; however, the impact seems to grow rapidly with increasing liability. The higher PRS-related risk among individuals with a family history suggests that the PRS likely consists of true causative alleles and may thus potentially shed light on the biological processes leading to schizophrenia. We attempted to assess whether family history was mediated through the PRS when the interaction was taken into account, while 17.4% (95% CI, 9.1-26.6) was mediated when the interaction was ignored. It is recognized that sporadic cases without a family history are common, although those with a family history may be expected to carry a higher genetic liability. Our results were
consistent with these expectations, implying that despite the correlational nature of the PRS, other pathways to schizophrenia may exist particularly among individuals without a family history.

Parental socioeconomic status was robustly associated with the risk for schizophrenia in offspring (OR range, 1.56; 95% CI, 1.18-2.05 to 8.10; 95% CI, 3.24-20.3). Our analyses suggest that 45.8% (95% CI, 36.1-55.5) of cases with schizophrenia could be attributed to low socioeconomic status. Schizophrenia may have insidious onset and may hinder social achievements long before the first episode, at which point the social decline tends to cease.4 The inverse association between the risk for schizophrenia and socioeconomic status has been ascribed to social selection (ie, that genetically predisposed persons drift downwards in terms of social class or fail to rise out of low social status26) but this has been questioned.27 We found no interaction, and our analyses did not suggest that socioeconomic status was mediated through the PRS; however, the wide confidence interval suggests there is a great deal of uncertainty about the proportion mediated. Because the parental genetic architecture was unobserved, this does not imply that socioeconomic status and genomic variants generally are unassociated. On the contrary, one study has reported a association with educational attainment.28

Our study had several limitations. First, schizophrenia is a heterogeneous disorder that varies across sex, age, race/ethnicity, and place. Although the discovery sample was historically large (34,600 cases and 45,968 control individuals),4 our results may not generalize because they were based on only 871 control individuals and 866 incident cases. Furthermore, criteria for diagnosing schizophrenia may vary between Denmark and other countries. Second, the PRS represents a mixture of true and false common schizophrenia risk alleles, where little is known about the biological pathway between any such allele and schizophrenia.16 Furthermore, the PRS captures only a proportion of the variation attributable to common SNPs20 and is not expected to capture deleterious exonic mutations or rare genetic and copy number variations that may be important for the development of schizophrenia. On the other hand, the PRS was obtained independently. Third, parental socioeconomic status at the individual's birth was based on observational data from registers. Although this minimizes differential misclassification, subsequent loss of income and job, marital breakup, or continued educational underachievement may correlate with the risk for schizophrenia. Furthermore, our measure of socioeconomic status was rather crude and may only be weakly related to the actual circumstances during an individual's upbringing. Nevertheless, each factor that constitutes socioeconomic status has both here and previously been associated with the risk for schizophrenia.27 Fourth, family history and diagnostic information relied on clinical diagnosis assigned by the attending physician at discharge; thus, phenotypic misclassification is an issue of concern. Our sample was too small to examine a more extended diagnostic classification. However, the Danish Psychiatric Research Register has high diagnostic validity,11 and because familial history is assessed independently of caseness and because the related rates were in keeping with previous findings,18 it is less likely that our sample was biased. Only information on maternal siblings was used. Fifth, the method of selecting at-risk control individuals and incident cases was unlikely to be optimal for identifying genetic markers, although this strategy was likely optimal for the family history and socioeconomic status criteria. Lastly, none of our risk factors were amenable to direct intervention; thus, our results were not directly translational but our analyses explored their mutual importance.

Conclusions

Schizophrenia was consistently related to the PRS, parental socioeconomic status, and family histories of psychiatric disorders. Despite their interdependencies, each factor accounted for a sizeable fraction of cases; however, only a modest part of the variation was explained. A considerable proportion of the association with family history of schizophrenia/psychoses was mediated through the PRS. A negligible fraction of parental socioeconomic status appeared to be mediated through the PRS.
Variables Associated With Schizophrenia Risk

support from H. Lundbeck A/S. No other disclosures were reported.

Funding/Support: This study was supported by grants from the Danish Strategic Research Council, the Faculty of Health Sciences at Aarhus University, the Lundbeck Foundation, the Stanley Medical Research Institute, and a European Research Council advanced grant (Dr Mortensen; GA 294838). Dr Wray acknowledges funding from the Australian National Health and Medical Research Council (grants 613602 and 1047956).

Role of the Funder/Sponsor: The funders had no role in the design of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES


