Title: Polygenic risk score for schizophrenia and treatment-resistant schizophrenia

Running title: Polygenic risk and treatment-resistant schizophrenia

Authors: Theresa Wimberley$^{1,2,3}$, Christiane Gasse$^{1,2,3}$, Sandra Melanie Meier$^{1,2,3,4,5}$, Esben Agerbo$^{1,2,3}$, James H MacCabe$^6$, Henriette Thisted Horsdal$^{1,2,3}$

1 National Centre of Register-based Research, School of Business and Social Sciences, Aarhus University
2 iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, DK
3 Centre for Integrated Register-based Research, CIRRAU, Aarhus University, Aarhus, Denmark
4 Child and Adolescent Mental Health Centre – Mental Health Services Capital Region, Copenhagen Region, DK
5 Psychosis Research Unit, Aarhus University Hospital, Risskov, DK
6 Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK

Corresponding author:
Theresa Wimberley,
National Centre for Register-based Research,
School of Business and Social Sciences, Aarhus University,
Fuglesangs allé 4, Building K,
DK-8210 Aarhus V,

Phone: +45 87165976,

e-mail: tw@econ.au.dk
**Abstract**

Treatment-resistant schizophrenia (TRS) affects around one third of individuals with schizophrenia. Although a number of socio-demographic and clinical predictors of TRS have been identified, data on the genetic risk of TRS are sparse. We aimed to investigate the association between the polygenic risk score for schizophrenia and treatment resistance in patients with schizophrenia. We conducted a nationwide, population-based follow-up study among all Danish individuals born after 1981 and with an incident diagnosis of schizophrenia between 1999 and 2007. Based on genome-wide data polygenic risk scores for schizophrenia were calculated in 862 individuals with schizophrenia. TRS was defined as either clozapine initiation or at least two periods of different antipsychotic monotherapies and still being hospitalized. We estimated hazard rate ratios (HRs) for TRS in relation to the polygenic risk score while adjusting for population stratification, age, sex, geographical area at birth, clinical treatment setting, psychiatric comorbidity, and calendar year. Among the 862 individuals with schizophrenia, 181 (21.0%) met criteria for TRS during 4,674 person-years of follow-up. We found no significant association between the polygenic risk score and TRS, adjusted HR = 1.13 (95% CI: 0.95-1.35). Based on these results, the use of the polygenic risk score for schizophrenia to identify individuals with TRS is at present inadequate to be of clinical utility at the individual patient level. Future research should include larger genetic samples in combination with non-genetic markers. Moreover, a TRS-specific developed polygenic risk score would be of great interest towards early prediction of TRS.

Keywords: genetic data, antipsychotics, clozapine, treatment resistance
Introduction

Treatment-resistant schizophrenia (TRS) is considered to affect approximately one third of all individuals diagnosed with schizophrenia. TRS is commonly defined as insufficient treatment response despite at least two different antipsychotic treatment courses of adequate dose and duration.\textsuperscript{1-3} It has been shown, that individuals with a family history of psychosis are more likely to poorly respond to antipsychotic drugs,\textsuperscript{4-6} suggesting a genetic influence in the development of TRS. This is supported by previous research indicating that variants in the gene coding for the Dopamine D2 receptor (DRD2) may be predictive of response to antipsychotic drugs in first-episode schizophrenia.\textsuperscript{7-9} The DRD2 lies within one of the 108 associated regions reported in schizophrenia showing an overlap between genes contributing to the genetic risk of schizophrenia and the key target for antipsychotic drugs.\textsuperscript{10} Schizophrenia is polygenic, \textit{i.e.} influenced by many genetic variants each having a small effect.\textsuperscript{11} Combining thousands of common variants of small effect, weighted by their effect estimates, in a polygenic risk score, provides a measure of genetic liability to schizophrenia.\textsuperscript{12} Data regarding the association of the polygenic risk score for schizophrenia and TRS are sparse,\textsuperscript{13,14} and more evidence is needed to explore this association further. As the social and economic costs of TRS are high,\textsuperscript{15} and an extended duration of insufficiently treated or untreated psychosis is associated with poorer prognosis,\textsuperscript{16} it is important to identify individuals at high risk of TRS to improve efficacy and reduce time and morbidity during failed treatments periods. Genetic predictors could potentially help to identify such high-risk individuals.

In this nationwide, population-based follow-up study, we aimed to examine the association between the polygenic risk score for schizophrenia and TRS in a Danish cohort of individuals diagnosed with schizophrenia.

Methods

This was a prospective cohort study using Danish, population-based registers.
**Data sources**

Danish, population-based registers were linked using the unique personal identification number assigned to all Danish citizens. Information on sex, date of birth, death, emigration, and current and past residence in Denmark was obtained from the Danish Civil Registration System established in 1968. Genetic data were extracted from dried blood spots taken in the first days after birth of nearly all infants born in Denmark since 1981 and stored in the Danish Newborn Screening Biobank. Information on antipsychotic medication was obtained from the Danish National Prescription Registry (DNPR) which holds information on all prescriptions redeemed from community pharmacies in Denmark since 1995. Data on psychiatric hospital diagnoses and admissions were extracted from the Psychiatric Central Research Register (PCRR) which contains information on all inpatient contacts since 1969 and outpatient contacts since 1995.

**Study cohort**

The study cohort consisted of all individuals with a first diagnosis of schizophrenia (the International Classification of Diseases, Tenth Revision, ICD-10 code F20) between 1999 and 2007, born in Denmark after May 1, 1981, with a DNA sample available from the Danish Newborn Screening Biobank.

**Treatment-resistant schizophrenia**

Our measure of TRS was based on data from the DNPR and PCRR. We defined TRS as first occurrence of either clozapine initiation or hospitalization due to schizophrenia during antipsychotic treatment within 18 months after at least two periods of different antipsychotic monotherapy lasting at least 6 weeks each. This treatment-based proxy, used in a recent published study, builds on international and Danish treatment guidelines and the Kane criteria.

**The polygenic risk score**

The polygenic risk score for schizophrenia was based on genome-wide data from the PGC discovery sample and data from the Danish Newborn Screening Biobank. We calculated polygenic risk scores based on the summary statistics (effect allele, effect size) derived from the discovery sample. The discovery sample
comprised 34,600 schizophrenia cases and 45,986 controls from the PGC Genome-Wide Association Study (GWAS) meta-analysis for schizophrenia, excluding the Danish cases. We selected SNPs associated at a p-value threshold of 0.05 or lower. This threshold was chosen in accordance with other studies including polygenic risk scores, and this less stringent threshold, compared with thresholds used in association analysis of single variants, is considered appropriate to achieve a balance between the number of false-positive and true-positive risk alleles. Thereby, we identified a total of 24,755 SNPs. In our target sample, our independent study population, the polygenic risk score was then calculated for each individual as the weighted sum of risk alleles at the selected SNPs with the weight being the effect estimates in the discovery sample.

The polygenic risk score, that is approximately normal distributed, was converted into a z-score by subtracting the mean and dividing by the standard deviation.

**Candidate predictors of TRS**

In one of the models, we included a selected subset of factors that have been shown to be associated with TRS. These factors were sex, age at first diagnosis of schizophrenia, geographical area at birth (born in the capital vs born outside the capital), psychiatric comorbidity (any inpatient psychiatric hospitalization in the year prior to first diagnosis of schizophrenia), and clinical treatment setting (in- or outpatient at first diagnosis of schizophrenia).

**Statistical analysis**

Individuals were followed from the date of their first hospital diagnosis of schizophrenia until TRS, death, emigration, or end of study (December 31, 2010), whichever came first. We estimated hazard rate ratios (HRs) for the association between the standardized polygenic risk score and the incidence of TRS, and adjusted for the candidate predictors listed above and calendar year. All models including the polygenic risk score for schizophrenia were additionally controlled for population stratification by including ten first genomic principal components.
We conducted several sensitivity analyses to examine the robustness of the results. First, due to the limited size of the cohort, we included the residual from a linear regression of the polygenic risk score on the candidate predictors instead of including the set of candidate predictors directly into the model. Second, we restricted the TRS definition to clozapine initiation only. Third, we restricted the cohort to adult individuals (≥18 years at first diagnosis of schizophrenia) since the antipsychotic treatment guidelines as well as clinical prescribing practice and response to antipsychotics differ between childhood- and adult-onset schizophrenia.31,32 Finally, all analyses were repeated with polygenic risk scores calculated based on SNPs selected using thresholds of p<0.01 and p<0.1, including 10,622 and 35,792 SNPs, respectively.

The proportional hazards assumption was checked by diagnostic plots. All statistical analyses were performed using SAS 9.3 and Stata 13.

**Results**

We identified 862 individuals diagnosed with schizophrenia with an available DNA sample. Median age at first schizophrenia diagnosis was 19 years (inter-quartile range: 17-21 years), 471 (54.6%) were males and 391 (45.4%) were females. A total of 181 (21.0%) fulfilled the TRS definition during follow-up. They were followed for a total of 4,674 person-years resulting in an incidence rate of TRS of 3.87 (95% CI: 3.31-4.44) per 100 person-years, and a median follow-up of 5.4 years (inter-quartile range: 4.1-7.1 years). Characteristics for these individuals according to TRS are presented in Table 1. The mean standardized polygenic risk score was nominally higher among individuals fulfilling the TRS definition (0.24) compared with individuals not fulfilling the TRS definition (0.20), but the difference was not statistically significant (p=0.58).

The adjusted hazard rate ratio for TRS associated with one standard deviation increase in the polygenic risk score was 1.13 (95% CI: 0.95-1.35). Only adjusting for population stratification, the estimate was down to 1.02 (95% CI: 0.88-1.17) (Table 2).
We also evaluated the HRs obtained for some of the known predictors for TRS in our analyses in order to qualify the validity of our dataset. In this cohort, the factors female sex, psychiatric comorbidity, and being inpatient at first diagnosis of schizophrenia were significantly associated with increased rate of TRS, whereas no significant associations were found for age at first schizophrenia diagnosis and capital birth (Table 2).

When including the residual of a linear regression of the polygenic risk score on the candidate predictors, the association between the polygenic risk score and TRS was virtually unchanged, adjusted HR=1.12 (95% CI: 0.93-1.33).

When restricting the TRS definition to clozapine initiation only, the association remained statistically insignificant, but the adjusted HR was slightly higher, HR=1.23 (0.97-1.56) (Supplementary Table 1).

We further examined the association restricted to adult individuals (≥18 years at first diagnosis of schizophrenia), although with a weaker statistical precision. Among this subset of individuals (n=584), we found no association between the polygenic risk score and TRS, adjusted HR=1.01 (0.80-1.27) (Supplementary Table 2).

Finally, we repeated the analysis with polygenic risk scores calculated based on SNPs selected using different thresholds, and found similar results. Adjusted HRs for TRS were 1.09 (95% CI: 0.92-1.29) for p<0.01 and 1.17 (95% CI: 0.98-1.41) for p<0.1.

**Discussion**

This study provided no statistically significant evidence for an association between an increased polygenic risk score for schizophrenia and TRS.

Our study is in agreement with two other studies. In a study from Australia by Martin and Mowry, no statistically significant association between the polygenic risk score and TRS was found, where TRS was defined as non-response to antipsychotics inspired by the modified Kane criteria. Similarly, in a study from Iceland by Sigurdsson and Ingimarsson, the polygenic risk score did not differ with regard to clozapine use (Abstract for the 23rd European Congress of Psychiatry: Polygenic Risk Scores and Metabolic Side
Effects in Clozapine Treatment of Schizophrenia, 2015;30(Supplement 1):93). Another study found crude statistically significantly increased polygenic risk score among people with schizophrenia treated with clozapine, but the association did not remain significant after adjustment for clinical variables. In our study, the HR estimate was insignificantly but nominally increased when lifetime clozapine treatment was used as a proxy for TRS. In summary, the evidence for an association between the polygenic risk score for schizophrenia and TRS is still sparse and not fully clarified.

The main strengths of the setting of this study are the uniformly organized Danish health care system allowing a population-based design and the ability to link different data sources with prospectively collected data independent of the study hypothesis and with virtually complete follow-up.

The validity of the schizophrenia diagnosis in the registers is high, but by only including people with schizophrenia with an available DNA sample, the sample was relatively young, and this may have introduced selection bias, thus limiting generalizability. The relatively young cohort might explain why we did not observe significant associations between TRS and age at first schizophrenia diagnosis and capital birth, as opposed to other studies, but in line with a study on early-onset schizophrenia and clozapine. Moreover, in accordance with studies of similar sample sizes, no association was found between TRS and sex and family history of schizophrenia.

Based on register data with lack of information on in-hospital drug treatment and clinical scores, TRS cannot be identified without misclassification, as discussed elsewhere. Still, we believe that this proxy definition of TRS, based not only on clozapine initiation but on history of different AP treatment is closer to the theoretical definition of TRS based on clinical guidelines which, in turn, are based on the modified Kane criteria. Nevertheless, sensitivity analysis restricting the TRS definition to clozapine initiation did not alter our conclusion.
The ability of our study to detect an association between the polygenic risk score for schizophrenia and TRS may also be hindered by the relatively small sample size. This small sample size also only allowed adjusting for few factors. The estimate was largely unaffected by adjustment for candidate predictors.

In spite of lack of association between the polygenic risk score and TRS, there might still be a genetic influence on TRS. An association between total copy number duplication burden genome wide and TRS has been found,\textsuperscript{14} and people with TRS have been shown to have an excess of rare disruptive variants in gene targets of antipsychotics.\textsuperscript{11} While genetic variations at the DRD2 have been associated with treatment response in schizophrenia,\textsuperscript{7-9} these variants alone only have a modest effect and are not clinical useful. The polygenic risk score thus remains a potentially useful tool for genetic risk models. Previously, the polygenic risk score for schizophrenia has been shown to be able to predict case-control status in independent samples, although the sensitivity and specificity of the polygenic risk score for schizophrenia were found to be low.\textsuperscript{10} The predictive accuracy of the scores may be optimized by accounting for the effects of linkage disequilibrium,\textsuperscript{34} and thus the scores may also improve the prediction of TRS.

**Conclusion**
The polygenic risk score for schizophrenia was not statistically significantly associated with TRS. This result suggests that the predictive power of this polygenic risk score in identifying TRS individuals is at present inadequate to be of clinical utility at the individual patient level. For future research on genetic markers for TRS and potential implementation in clinical practice, larger genetic samples and possible combination with non-genetic markers will be required. Moreover, development and validation of a TRS-specific polygenic risk score would be of great importance.

**Supplemental material**
Supplemental material is available.

**Funding**
This work was supported by the European Community’s Seventh Framework Programme (FP7/2007-2013) (grant number 279227); and the Lundbeck Foundation, Denmark (grant number R155-2014-1724).
Dr MacCabe was funded by the NIHR Biomedical Research Centre for Mental Health at King’s College London and the South London and Maudsley NHS Foundation Trust. Dr Meier received further funding from the Mental Health Services, Capital Region, Copenhagen, Denmark. All other authors have declared that there are no conflicts of interest in relation to the subject of this study.

Acknowledgements

This study was approved by the Danish Data Protection Agency and the Danish Health and Medicines authority.


### Tables

Table 1: Characteristics for individuals with schizophrenia with a DNA sample available (N=862).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>TRS during follow-up</th>
<th>No TRS during follow-up</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>181</td>
<td>681</td>
<td>862</td>
</tr>
<tr>
<td>Polygenic risk score (mean, SD)</td>
<td>0.24 (1.04)</td>
<td>0.20 (1.01)</td>
<td>0.21 (1.02)</td>
</tr>
<tr>
<td>Age at first schizophrenia diagnosis, years (median, inter-quartile range)</td>
<td>19.0 (16.9-20.8)</td>
<td>19.2 (17.0-20.9)</td>
<td>19.1 (17.0-20.9)</td>
</tr>
<tr>
<td>Sex, N (%)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>101 (55.8%)</td>
<td>290 (42.6%)</td>
<td>391 (45.4%)</td>
</tr>
<tr>
<td>Male</td>
<td>80 (44.2%)</td>
<td>391 (57.4%)</td>
<td>471 (54.6%)</td>
</tr>
<tr>
<td>Geographical area at birth, N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital</td>
<td>25 (13.8%)</td>
<td>116 (17.0%)</td>
<td>141 (16.4%)</td>
</tr>
<tr>
<td>Provincial or rural</td>
<td>156 (86.2%)</td>
<td>565 (83.0%)</td>
<td>721 (83.6%)</td>
</tr>
<tr>
<td>Treatment setting at first diagnosis of schizophrenia, N (%)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inpatient</td>
<td>121 (66.9%)</td>
<td>306 (44.9%)</td>
<td>427 (49.5%)</td>
</tr>
<tr>
<td>Outpatient</td>
<td>60 (33.1%)</td>
<td>375 (55.1%)</td>
<td>435 (50.5%)</td>
</tr>
<tr>
<td>Psychiatric comorbidity in the year prior to first diagnosis of schizophrenia, N (%)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>91 (50.3%)</td>
<td>208 (30.5%)</td>
<td>299 (34.7%)</td>
</tr>
<tr>
<td>No</td>
<td>90 (49.7%)</td>
<td>473 (69.5%)</td>
<td>563 (65.3%)</td>
</tr>
</tbody>
</table>

* p<0.05 for significant difference between individuals with treatment-resistant and non-treatment-resistant schizophrenia.
Table 2: Hazard rate ratios (HR) for the event of TRS (crude and adjusted associations with TRS) (N=862).

<table>
<thead>
<tr>
<th></th>
<th>Crude HR (95% CI)</th>
<th>Adjusted HR (95% CI)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polygenic risk score</td>
<td>1.02 (0.88-1.17)²</td>
<td>1.13 (0.95-1.35)²</td>
</tr>
<tr>
<td>Age at first diagnosis of schizophrenia (years)</td>
<td>1.02 (0.97-1.08)</td>
<td>1.02 (0.96-1.08)</td>
</tr>
<tr>
<td>Female</td>
<td>1.68 (1.25-2.25)</td>
<td>1.57 (1.16-2.13)</td>
</tr>
<tr>
<td>Capital birth</td>
<td>0.74 (0.49-1.13)</td>
<td>0.76 (0.48-1.19)</td>
</tr>
<tr>
<td>Inpatient at first diagnosis of schizophrenia</td>
<td>2.15 (1.58-2.93)</td>
<td>2.07 (1.50-2.84)</td>
</tr>
<tr>
<td>Any psychiatric comorbidity in the year prior to first diagnosis of schizophrenia</td>
<td>2.10 (1.57-2.81)</td>
<td>1.86 (1.38-2.51)</td>
</tr>
<tr>
<td>Calendar year</td>
<td>1.02 (0.94-1.10)</td>
<td>1.04 (0.95-1.13)</td>
</tr>
</tbody>
</table>

¹ Adjusted for age, sex, geographical area at birth, treatment setting, psychiatric comorbidity, and calendar year.
² Including the first ten genomic principal components to adjust for population stratification.
### Supplementary Table 1: Hazard rate ratios (HR) for the event of clozapine initiation (crude and adjusted associations with clozapine initiation) (N=862).

<table>
<thead>
<tr>
<th></th>
<th>Crude HR (95% CI)</th>
<th>Adjusted HR (95% CI)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polygenic risk score</td>
<td>1.08 (0.90-1.30) ²</td>
<td>1.23 (0.97-1.56) ²</td>
</tr>
<tr>
<td>Age at first diagnosis of schizophrenia (years)</td>
<td>1.06 (0.99-1.13)</td>
<td>1.09 (1.01-1.19)</td>
</tr>
<tr>
<td>Female</td>
<td>1.39 (0.95-2.05)</td>
<td>1.35 (0.91-2.00)</td>
</tr>
<tr>
<td>Capital birth</td>
<td>0.74 (0.43-1.30)</td>
<td>0.81 (0.45-1.46)</td>
</tr>
<tr>
<td>Inpatient at first diagnosis of schizophrenia</td>
<td>2.17 (1.44-3.28)</td>
<td>2.02 (1.32-3.09)</td>
</tr>
<tr>
<td>Any psychiatric comorbidity in the year prior to first</td>
<td>2.28 (1.55-3.35)</td>
<td>1.98 (1.33-2.94)</td>
</tr>
<tr>
<td>diagnosis of schizophrenia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calendar year</td>
<td>0.96 (0.87-1.05)</td>
<td>0.94 (0.84-1.06)</td>
</tr>
</tbody>
</table>

¹ Adjusted for age, sex, geographical area at birth, treatment setting, psychiatric comorbidity, and calendar year.

² Including the first ten genomic principal components to adjust for population stratification.

### Supplementary Table 2: Hazard rate ratios (HR) for the event of TRS among adult individuals (crude and adjusted associations with TRS) (N=584).

<table>
<thead>
<tr>
<th></th>
<th>Crude HR (95% CI)</th>
<th>Adjusted HR (95% CI)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polygenic risk score²</td>
<td>0.92 (0.75-1.11) ²</td>
<td>1.01 (0.80-1.27) ²</td>
</tr>
<tr>
<td>Age at first diagnosis of schizophrenia (years)</td>
<td>1.02 (0.91-1.14)</td>
<td>1.07 (0.93-1.22)</td>
</tr>
<tr>
<td>Female</td>
<td>1.28 (0.87-1.86)</td>
<td>1.08 (0.72-1.61)</td>
</tr>
<tr>
<td>Capital birth</td>
<td>0.59 (0.31-1.13)</td>
<td>0.65 (0.34-1.27)</td>
</tr>
<tr>
<td>Inpatient at first diagnosis of schizophrenia</td>
<td>2.29 (1.54-3.40)</td>
<td>2.30 (1.53-3.46)</td>
</tr>
<tr>
<td>Any psychiatric comorbidity in the year prior to first</td>
<td>1.92 (1.32-2.80)</td>
<td>1.74 (1.18-2.58)</td>
</tr>
<tr>
<td>diagnosis of schizophrenia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calendar year</td>
<td>0.98 (0.88-1.09)</td>
<td>0.99 (0.88-1.13)</td>
</tr>
</tbody>
</table>

¹ Adjusted for age, sex, geographical area at birth, treatment setting, psychiatric comorbidity, and calendar year.

² Including the first ten genomic principal components to adjust for population stratification.