Clinical and genetic predictors of treatment resistant psychosis

Smart, Sophie

Awarding institution:
King's College London

Clinical and Genetic Predictors of Treatment Resistant Psychosis

Sophie Elizabeth Smart

Thesis submitted for the degree of Doctor *of* Philosophy

Institute of Psychiatry, Psychology & Neuroscience, King's College London

2019

1. Acknowledgements

First and foremost, I would like to thank my supervisors. To Professor James MacCabe, thank you for giving me countless opportunities throughout this PhD and having endless faith in my ability to get the job done. To Professor Robin Murray, for always showing me the bigger picture and, when it was most needed, smoothing the road ahead. To Dr Paul O'Reilly, for allowing me to think out loud and teaching me to think that little bit further. And to Professor James Walters, my unofficial fourth supervisor, thank you for your help and encouragement over all these years.

This PhD would not have been possible, nor as enjoyable, without Dr Antonio Pardiñas and Dr Deborah Agbedjro. I could not have asked for two better teachers; I am sincerely grateful for your time, patience, enthusiasm, knowledge, encouragement, and support. Antonio – your explanations of complicated concepts in statistical genetics over email have been invaluable (I have saved them all!) and I am eternally grateful that I could include your work in this thesis. Deborah – your ability to problem solve is phenomenal, I am so lucky to have had the opportunity to work with you. Thank you, both, for your kindness and generosity.

The data for this PhD only came together because of a lot of hard work (and emails) from researchers all over the world. It has been a pleasure working with you. I also know that a lot of people have worked on these studies, long before this PhD even started, and am equally appreciative of their work. To all the participants in STRATA-G, I only met a few of you, but I am enormously grateful that you agreed to take part in your respective studies.

I need to thank three incredible people who have, in different ways, championed me throughout this PhD: Laura Kassoumeri, Dr Gemma Modinos, and Dr Sarah Davidson. Thank you for your friendship and support.

I would not have been able to complete this PhD without my life outside it. To all my friends and colleagues in PhD-land at the IoPPN, but especially Fiona, Alan, Amy, Cathy, Lilla, George, Sandra, Uzma – thank you for your camaraderie. It has

been a joy to work with everyone on STRATA, especially my fellow recruiters on STRATA 2. Special thanks must go to Kira Griffiths who proof-read parts of this thesis. To all my family and friends outside of the IoPPN, you make life so enjoyable. Thank you to Rachael who, long ago, when we were small, believed that I could write. And thank you to Nicole and Elisha for being there throughout all the highs and lows. Elisha – I wanted to make this thesis as readable as possible for non-scientists and, if I have in any way accomplished this aim, it is in part because of your thoughtful comments and careful edits.

Finally, and most importantly, I need to thank my family: Mum and Dad, Patrick and Cameron. You are my home, no matter how far apart we all are, and everything I have achieved has only been possible because of you.

Sophie Elizabeth Smart

October 2019

2. Preface

2.1. Statement of Personal Contribution

For this thesis, I organised data collation of the STRATA-G samples and conducted all data management and cleaning. I also conducted additional follow-up recruitment for the AESOP study.

I wrote this thesis in its entirety, with the following exceptions:

- Chapter 8 (systematic review), after being drafted by me, was circulated to co-authors and underwent peer review prior to acceptance in Psychological Medicine.
- Chapter 9 (prediction model), after being drafted by me, was circulated to co-authors who provided comments.
- Chapter 10 (genome-wide association study), is my summary of the work led by Dr Antonio Pardiñas. A draft manuscript detailing this work, written by Dr Antonio Pardiñas and commented on by me, is reproduced in Appendix D with his permission.

I designed and executed all the analyses presented in Chapter 9 (prediction model) and had the assistance of Dr Deborah Agbedjro in writing analysis scripts. For Dr Pardiñas' work presented in Chapter 10 (genome-wide association study), I collated the STRATA-G sample used in the polygenic risk score analysis and implemented the quality control pipeline on the genotypes for STRATA-G. I designed and performed all the analyses presented in Chapter 11 (path analysis), using the polygenic risk scores calculated by Dr Pardiñas.

2.2. Thesis incorporating publications

This thesis is a "thesis incorporating publications". This refers to the fact that one chapter is composed of a published journal article of which I am the first author.

Chapter 8 is composed of the following journal article which is reproduced in full:

Smart, S. E., Kepinska, A. P., Murray, R. M. & MacCabe, J. H. (2019). Predictors of treatment resistant schizophrenia: a systematic review of prospective observational studies. *Psychological Medicine*, e-pub ahead of print.

The final published version of this article is presented in Appendix B.

2.3. Terminology: psychosis versus schizophrenia

I have not limited this thesis to treatment resistant schizophrenia and have instead chosen to focus on treatment resistant psychosis. There is a lack of evidence to suggest that schizophrenia is a distinct and stable entity (Guloksuz & Van Os, 2018). Antipsychotics primarily reduce psychotic symptoms, and there is limited evidence that they improve the other symptoms of schizophrenia: negative symptoms (Harvey, James, & Shields, 2016; Krause et al., 2018) and impaired cognition (Takeuchi, Thiyanavadivel, Fervaha, & Remington, 2017). By examining the clinical and genetic predictors of treatment resistance in people with a broad range of psychosis spectrum disorders, I hope the results of this thesis will be generalisable to more patients.

In Chapter 8, the terminology 'treatment resistant schizophrenia' (TRS) is used to be consistent with the published version of the systematic review. However, this review included patients with a range of psychosis spectrum disorders and is therefore comparable to the rest of this thesis.

In Appendix D, TRS is used in the draft manuscript written by Dr Pardiñas, although the STRATA-G sample included in this work included patients with a range of psychosis spectrum disorders. In all other Chapters, the terminology 'treatment resistance' or 'treatment resistance psychosis' is used and both are abbreviated to TR.

3. Abstract

Treatment resistance (TR), affecting approximately 20-30% of patients with psychosis, has a high burden both for patients and healthcare services. I provided a general introduction to TR in Chapter 6 and highlighted the need to identify TR earlier in the course of the illness, so that an effective treatment, such as clozapine, can be offered promptly. The purpose of this thesis was to identify clinical, demographic, and genetic predictors of TR, using information identified when a patient first presents to clinical services with psychosis.

First, I reported the results of a systemic literature review, which synthesised predictors of TR identified in longitudinal, observational studies (Chapter 8). Younger age of onset was the most consistent predictor of TR, but this review also indicated that, to date, studies have not used statistical methods specifically designed to identify predictors that have a high chance of predicting TR in future patients. Existing literature has primarily used methods designed to capture the magnitude of association, between predictors and TR, within the study sample.

Second, I reported the results of my own analysis using a dataset created by combining existing longitudinal, first episode psychosis cohorts. This dataset, known as STRATA-G (and described in Chapter 7), included patients who had a minimum of one year follow-up, provided a DNA sample, could be classified as either TR or non-treatment resistant (NTR). I found significant associations between TR and both younger age of onset and fewer years in education (Chapter 9). I created a predictive model of TR that selected nine first episode variables and could, after internal validation, correctly classify 59% of TR patients and 65% of NTR patients (Chapter 9).

Third, I reported the results of a genome-wide association study (GWAS) of treatment resistance, which was used to create a polygenic risk score for TR (Chapter 10). This polygenic risk score was significantly associated with TR in the STRATA-G sample.

Fourth, I investigated whether age of psychosis onset mediated the relationship between genetic burden for TR and subsequent TR, using STRATA-G (Chapter 11). Age of onset was not associated with genetic risk for TR and there was no evidence it mediated the relationship between the polygenic risk score for TR and subsequent TR.

Finally, I expanded on the discussion of the findings in the preceding chapters (Chapter 12). Here, I placed my findings in the context of the wider literature and discussed the limitations of my data. I discussed the future of prediction modelling in TR; the potential uses of a prediction model, in both clinical practice and research studies, and the importance of future work to validate the findings reporting in this thesis.

4. Table of Contents

Table of Tables

Table of Figures

5. List of Abbreviations

95%CI	95% confidence intervals
A level	General Certificate of Education (GCE) Advanced Level
AESOP	Aetiology and Ethnicity in Schizophrenia and Other Psychoses study
AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
AUC	Area under the receiver operating characteristic curve
Baccalaureate	An alternative to A Levels
BMI	Body Mass Index
BoFEP	Bologna first episode psychosis study
BPRS	Brief Psychiatric Rating Scale
BTEC	Business and Technology Education Council level 3 vocational qualifications
CAARMS	Comprehensive Assessment of At Risk Mental States scale
CAG	Confidentiality Advisory Group
CardiffCOGS	Cross-sectional psychosis study
CATIE	Clinical Antipsychotic Trials of Intervention Effectiveness study
CBT	Cognitive behavioural therapy
CFI	Comparative Fit Index
CGI	Clinical Global Impression scale
CI	Chief Investigator
CLOZ-PRS	Polygenic risk score for clozapine use
CLOZUK1	A sample of samples of TRS individuals
CLOZUK2	A sample of samples of TRS individuals
CMRS	Case Manager Rating Scale
CNIL	National Commission on Informatics and Liberty
CRP	C-reactive protein
CU	Cardiff University
DIP	Diagnostic Interview for Psychosis
DNA	Deoxyribonucleic acid
DSM	Diagnostic Statistical Manual
DTA	Data transfer agreements
DUP	Duration of untreated psychosis
EA-PRS	Polygenic risk score for educational attainment
EUGEI	European Network of National Schizophrenia Networks Studying Gene-Environment Interactions study
FDA	Food and Drug Administration
FDR	False discovery rate
FGA	First-generation antipsychotics
FIGS	Family Interview for Genetic Studies
GAF	Global Assessment of Functioning
GAF-F	Global Assessment of Functioning – functioning scale
GAF-S	Global Assessment of Functioning – symptom scale
GAP	Genetics and Psychosis study
GCSE	General Certificate of Secondary Education
GWAS	Genome-wide association studies
HR	Hazard ratio
HRA	Health Research Authority
HWE	Hardy–Weinberg equilibrium
IBS	Identity-by-state

ICD	International Classification of Diseases
ICD	International Classification of Diseases
IoPPN	Institute of Psychiatry, Psychology & Neuroscience
IQ	Intelligence quotient
IQ-PRS	Polygenic risk score for IQ
IQR	Inter quartile range
INFO	Measure of the quality of imputation
KCL	King's College London
LD	Linkage disequilibrium
MAF	Minor allele frequency
MRC	Medical Research Council
MRI	Magnetic resonance imaging
MTA	Material transfer agreements
NHST	Null hypothesis significance test
NICE	National Institute for Health and Care Excellence
NIFEPS	Northern Ireland First Episode Psychosis study
NTR	Non-treatment resistance
NVQ	National Vocational Qualification
O levels	General Certificate of Education (GCE) Ordinary Level
OPCRIT	Operational Criteria Checklist for Psychotic Illness and Affective Illness
OR	Odds ratio
PAFIP	First Episode Psychosis Clinical Program study
PANSS	Positive and Negative Syndrome Scale
PAS	Premorbid Assessment Scale
PC/PCs	Principal components
PET	Positron-emission tomography
PGC	Psychiatric Genomics Consortium
PGC-PRS	Polygenic risk score for non-treatment resistance schizophrenia
PI	Principal Investigator
PRS	Polygenic risk score
REC	Research Ethics Committee
RGPI	Resources for Genomics, Ireland study
RMSEA	Root Mean Square Error of Approximation
SANS	Scale for the Assessment of Negative Symptoms
SAPS	Scale for the Assessment of Positive Symptoms
SCAN	Schedules for Clinical Assessment in Neuropsychiatry
SGA	Second-generation antipsychotics
SNP	Single nucleotide polymorphism
SOFAS	Social and Occupational Functioning Assessment Scale
STRATA	Schizophrenia: Treatment Resistance and Therapeutic Advances
STRATA-G	STRATA-Genetics
SZ-PRS	Polygenic risk score for schizophrenia
TIPP	Treatment and Early Intervention in Psychosis Program study
TOP	Thematic Organized Psychosis Research study
TR	Treatment resistant psychosis
TR-PRS	Polygenic risk score for treatment resistance
TRRIP	Treatment Response and Resistance in Psychosis
TRS	Treatment resistance schizophrenia
UCL	University College London
WS3	Workstream three of STRATA

INTRODUCTION

6. Introduction and Aims

6.1. Psychosis

The term 'psychosis' refers to a heterogenous cluster of symptoms that commonly occur together in schizophrenia and related psychosis spectrum disorders (Carpenter Jr, 2007; Kahn et al., 2015; Tamminga & Holcomb, 2005). Psychosis is the distortion of reality where an individual can experience grandiosity, paranoia, hallucinations, and unusual thought content. In psychosis spectrum disorders, psychosis is often accompanied by so-called negative symptoms (avolition, apathy, anhedonia, blunted affect), disorganisation (conceptual disorganisation, disorientation, posturing), cognitive impairment (poor memory, attention, and fluency), and affective symptoms (depression, anxiety, guilt). Two patients may meet the diagnostic criteria for the same disorder, for example schizophrenia, but experience different symptoms, have different underlying aetiologies, and respond differently to treatment.

Some claim that psychotic experiences can occur in as many as 7.8% of the global population (McGrath et al., 2016) but the vast majority of these people never need to seek treatment. On the other hand, 1 in 200 individuals will specifically develop schizophrenia in their lifetime (median lifetime prevalence of 0.48%; Simeone, Ward, Rotella, Collins, & Windisch, 2015) and 1.5 in 200 individuals will develop a disorder that falls within the broad range of psychosis spectrum disorders (median lifetime prevalence of 0.75%; Moreno-Küstner, Martín, & Pastor, 2018). As such, psychotic disorders are among the leading cause of disability across the world (Vos et al., 2015) with those affected experiencing a marked decrease in quality of life (Bobes, Garcia-Portilla, Bascaran, Saiz, & Bousoño, 2007). 10-20% of patients experience just one episode of psychosis and spend the rest of their lives in recovery without long-term functional or social impairments. For other patients, the illness course consists of multiple psychotic episodes separated by periods of symptomatic, although not always functional, remission (Carbon & Correll, 2014; Emsley, Chiliza, Asmal, & Harvey, 2013; Häfner, 2014). Although this suggests that psychosis spectrum disorders are chronic and lifelong conditions, many people experience symptomatic remission and recovery. One study, which followed a cohort of first episode psychosis patients for ten years, reported that 65% of patients

were in remission at the 10 year point and 46% of patients had been symptom free for the preceding two years (and thus met criteria for symptomatic recovery; Revier et al., 2015). Of those in recovery, 56% of patients has been prescribed antipsychotic medication in the preceding two years.

6.2. Antipsychotics

Antipsychotic medication is the most widely used – and offered – treatment for psychosis (Carter et al., 2018). Antipsychotics are efficacious (Leucht et al., 2013) and, when compared to placebo, reduce relapse rates and increase patients' quality of life (Leucht et al., 2012). A recent meta-analysis pooled response rates across antipsychotic trials: 81% of patients experienced a 20% (or more) reduction in symptom severity and 52% experienced a 50% (or more) reduction (Zhu et al., 2017). Antipsychotics are also highly effective; in a sample of 311 patients, 79% reported that antipsychotics had made them feel better (Carter et al., 2018). The introduction of antipsychotics has greatly improved the lives of many people with psychosis. However, it is important to remember that – despite the focus of this thesis on antipsychotic treatment resistance – some patients experience remission without antipsychotic use (Wils et al., 2017).

Antipsychotics can be split into two categories: first-generation antipsychotics (FGAs), which were developed in the 1950s, and second-generation antipsychotics (SGAs) that have been available since the 1990s. There is considerable controversy over whether any one antipsychotic, or one generation, is superior to the others. Meta-analyses tend to favour only four SGAs – amisulpride, olanzapine, risperidone and clozapine (more on clozapine below) – over FGAs, but overall SGAs are associated with fewer extrapyramidal side effects and possibly less negative symptoms than FGAs (Fusar-Poli et al., 2015; Leucht et al., 2013; Leucht et al., 2009). In terms of treatment management, long-acting injectables are superior to oral formulations in preventing rehospitalisation (Tiihonen et al., 2017b), not because of any greater efficacy but simply because the patient is more likely to actually receive the drug. Maintenance treatment for at least three years, for those who fully responded to medication, decreases the risk of relapse (Hui et al., 2018).

The mechanism of action of antipsychotics has been used to reverse-engineer a 'dopamine hypothesis' of psychosis (Howes & Kapur, 2009). In brief, patients with psychosis have increased presynaptic dopamine synthesis capacity and increased levels of synaptic dopamine; antipsychotics bind to dopamine D2 receptors in the brain; and higher D2 receptor occupancy is associated with a greater reduction in psychotic symptoms (Howes & Murray, 2014; Kapur, Agid, Mizrahi, & Li, 2006).

Despite the success of antipsychotics in treating psychosis, they are not a panacea for all patients. Every single antipsychotic medication comes with the risk of at least one side effect (Leucht et al., 2013). Antipsychotic side effects can be severe, intolerable, and in some cases life-threatening. They include involuntary motor movements (extrapyramidal side effects), weight gain and associated metabolic problems including diabetes and cardiac arrest, sedation, QTc prolongation leading to torsades de pointes, and increases in prolactin. Even if side effects are tolerable, it is not guaranteed that symptoms will respond to treatment. In the large CATIE study, 74% of patients discontinued or switched antipsychotic medication in the first eighteen months; 15% discontinued or switched owing to the intolerability of side effects but 24% of patients discontinued or switched antipsychotic medication due to a lack of efficacy (Lieberman et al., 2005). It is these patients – treated with antipsychotics but still experiencing psychotic symptoms – that are the focus of this thesis.

6.3. Treatment resistant psychosis

In the UK, the definition of treatment resistance (TR) has been standardised as the failure to respond to two sequential trials of *different* antipsychotics, where a trial is defined as a period of 4-6 weeks during which antipsychotics are prescribed at the optimum dosage (NICE guidelines; National Institute for Health and Care Excellence, 2014). TR is estimated to occur in 20-30% of patients (Elkis & Buckley, 2016). In an influential open-label trial, 75% of patients achieved remission after one trial, a further 4% remitted after a second, which left a final 21% who met criteria for TR (Agid et al., 2011). The prevalence of TR varies depending on the clinical setting and stage of illness, as such, estimates are higher in community-based studies (56%; Beck et al., 2019).

Although many TR criteria specifically exclude intolerance, including the NICE guidelines, some early studies classified both patients who are resistant to antipsychotics and patients who are intolerant to their side effects as TR. This is because the treatment for TR, clozapine, was originally licensed by the Food and Drug Administration (FDA) in the USA for both lack of response and intolerance (Conley & Kelly, 2001). Of course, in all studies there is a risk that patients will be incorrectly classified as TR; including patients who are intolerant to side effects, nonadherent to treatment, and those for whom the recommended antipsychotic dose is subtherapeutic. The limitations of how TR is defined are discussed in depth in Chapter 12.

Persistent psychotic symptoms, as is the case in TR, have adverse consequences for the health and well-being of patients and their families. Patients have a poor quality of life, poor physical health, and their level of disability prevents them from working (Brain, Kymes, DiBenedetti, Brevig, & Velligan, 2018; Iasevoli et al., 2016; Nordstroem, Talbot, Bernasconi, Berardo, & Lalonde, 2017). This, in combination with the cost of medical treatment and hospitalisation, increases the societal economic burden of TR (Andrews, Knapp, McCrone, Parsonage, & Trachtenberg, 2012; Kennedy, Altar, Taylor, Degtiar, & Hornberger, 2014). One study estimated that the health costs associated with TR were 3 to 11-fold higher than non-treatment resistant psychosis (Kennedy et al., 2014)

The aetiology of TR is as yet unknown. Some believe that TR simply encapsulates the severe end of psychosis while others consider that TR is biologically distinct from NTR. However, TR does appear often to be a trait, rather than a state, since 70-84% of patients meet TR criteria without ever experiencing a period of remission (Demjaha et al., 2017; Lally et al., 2016a).

Despite a lack of understanding of the cause(s) of TR, TR itself is treatable. There is evidence that cognitive behavioural therapy (CBT) as adjunct treatment may be effective (Morrison et al., 2018) but it is clozapine which is synonymous with treatment for TR. Clozapine, while technically an SGA, is distinct from other antipsychotics because it may only be used as a third-line treatment for psychosis; to be prescribed clozapine, patients must meet criteria for TR. Clozapine treatment

is accompanied by weekly blood monitoring to prevent some of its more severe side effects, particularly agranulocytosis. As clozapine is only licensed for TR, unless in very rare indications, its use is often used in retrospective and case registry studies as a definition of TR. While for a small proportion of patients, referred to as ultra-TR, clozapine will not be effective, it is highly effective for TR in terms of reducing symptom severity and mortality rates (Siskind, Siskind, & Kisely, 2017; Warnez & Alessi-Severini, 2014).

A significant problem in treating TR is the delay in clozapine prescription (Warnez & Alessi-Severini, 2014), with one study estimating that clozapine is delayed by an average of four years (Howes et al., 2012). In fact, delay in successful treatment throughout the course of psychosis is problematic. It is well known that patients with a longer delay between the time of their first psychotic symptom and the onset of treatment (duration of untreated psychosis) are less likely to respond to antipsychotics (Murru & Carpiniello, 2018). But it is also true that patients with a longer delay before successful treatment (duration of illness) are less likely to respond to antipsychotics. In a study by Malla and colleagues (Malla et al., 2006), patients with a longer duration of illness (defined as the time between a patient's first psychiatric symptom and adequate treatment, with adequate treatment being defined as four weeks of antipsychotic treatment or a significant response to treatment, whichever came first) had a poorer response to clozapine than those with a short duration. In addition, poor response in the first four weeks of antipsychotic treatment is a robust predictor of a lack of response later in the course of the illness (Carbon & Correll, 2014). Finally, a longer delay in clozapine initiation is associated with a poorer symptomatic response to clozapine (Yoshimura, Yada, So, Takaki, & Yamada, 2017).

The above provides a compelling case for attempting to identify TR earlier in the course of illness, and as such is the focus of this thesis. Early identification – or prediction – would not only be clinically beneficial for patients, but also useful in terms of research. The ability to identify patients at risk of being TR would create a new pool of research participants. Biological studies investigating the aetiology of TR, and clinical trials for current and novel compounds to treat TR, could recruit

patients much earlier in the course of their illness and prior to antipsychotic treatment, a potential confounder in such studies.

There are numerous strategies for early identification, both in terms of the data to be used and statistical methodology applied, but this thesis will focus solely on the use of clinical and genetic data. Clinical data is cheap and easy to collect, particularly when compared to neuroimaging data, and is non-invasive for patients. Genetic data *is* invasive but requires only a small blood draw and, with the proliferation of automated high-throughput array-based methods for genotyping, the cost of genotyping one individual is now ~£50. Below, I summarise the existing literature on the clinical phenotype and genetics of TR.

6.4. The clinical phenotype of treatment resistance

Many studies have compared TR to non-treatment resistant (NTR) patients, with the aim of identifying the phenotype which is unique to TR. Gillespie, Samanaite, Mill, Egerton, and MacCabe (2017), for example, published a comprehensive review in this field.

Naturally, patients with TR tend to have more severe symptoms than NTR patients because they are not receiving successful treatment. When using the Positive and Negative Syndrome Scale (PANSS), patients with TR have a higher total score, but also score higher on the negative symptoms subscale (Iasevoli et al., 2018a). This is true not only of TR patients with a diagnosis of schizophrenia but also those who have bipolar disorder or a psychotic disorder with comorbid depression or anxiety (Iasevoli et al., 2016). Severe psychotic and negative symptoms both have an adverse effect on quality of life (Narvaez, Twamley, McKibbin, Heaton, & Patterson, 2008; Watson et al., 2018). Yet, other research suggests that the symptoms structure within TR and NTR is remarkably similar. Freitas et al. (2019) tested six established factor models created using the PANSS, as well as their own factor model, yet no model was able to discriminate between TR and NTR.

In terms of cognition, TR patients are more severely impaired, particularly in terms of verbal memory, (de Bartolomeis et al., 2013; Frydecka, Beszłej, Gościmski, Kiejna, & Misiak, 2016; Joober et al., 2002). There is also evidence to suggest that

the cognitive impairments associated with TR actually precede the onset of psychosis, with TR patients having a lower premorbid IQ (Legge et al., 2019). One study however found no differences in cognitive impairment (Anderson, McIlwain, Kydd, & Russell, 2015). Impairment is not limited to cognitive abilities, with TR patients exhibiting worse social impairment too (Iasevoli et al., 2016). It can be presumed that impairments in cognitive and social functioning will have an adverse impact on a patient's ability to work and on their personal relationships. Indeed, patients with TR are less likely to be employed and less likely to be in full-time employment, although the proportion of TR patients in a long-term relationship or ever having been married is no different to the proportion of NTR (Iasevoli et al., 2016). Importantly, this study ruled out pseudo-TR by incorporating two prospective antipsychotic trials into their study design. Like cognitive impairment, functional impairment in work and social domains occurs in the premorbid period, and the impairment is more severe for patients with TR compared to NTR (Legge et al., 2019). Kravariti et al. (2018) measured cognition at first episode, and then ten years later compared patients who met criteria for TR to patients who had responded to antipsychotic treatment. They found that TR patients had poorer verbal intelligence and verbal fluency. Psychopathology and cognitive ability often interact, and, within TR, negative symptoms appear to mediate the effect of verbal fluency and neurological soft signs (often increased in patients with TR) on symptom severity (Iasevoli et al., 2018b). Impaired cognition and more severe negative symptoms also contribute to the functional impairment seen in TR patients (Iasevoli et al., 2018c).

In terms of other illness-related features, patients with TR consistently have a younger age of psychosis onset than their NTR counterparts (Demjaha et al., 2017; Lally et al., 2016a). The ability of age of onset, on its own, to discriminate between TR and NTR has been tested in two studies. In Iasevoli et al. (2018a), age of onset could not discriminate between TR and NTR, but in a study by Legge et al. (2019) younger age of onset could predict TR and explained 7.3% of the variance. There is tentative evidence that men are more likely to be TR (Lally et al., 2016a; Meltzer et al., 1997), but this may be confounded by age of onset since men have on average a younger age of onset than woman (Meltzer et al., 1997). Comorbidities are also common in TR; patients with TR are more likely to have a comorbid personality

disorder (Wimberley et al., 2016b), obsessive compulsive disorder or symptoms (Cunill, Castells, & Simeon, 2009; Sa et al., 2009), and autism spectrum disorders (Downs et al., 2017). TR is also associated with a higher risk of suicide (Kennedy et al., 2014; Wimberley et al., 2016b), experience of four or more stressful life events in childhood (Hassan & De Luca, 2015), and childhood sexual abuse (Hassan & De Luca, 2015), although not childhood abuse in general (Legge et al., 2019).

Patients with TR are more likely to smoke cannabis and to have used cannabis for longer than NTR patients (Arsalan et al., 2019). This cannot be easily dismissed as self-medicating in response to a more severe psychopathology as patients with TR are also more likely to use cannabis in the year prior to psychosis onset (Legge et al., 2019). Smoking tobacco is, however, not associated with TR, despite evidence that people with schizophrenia who smoke have worse symptoms, poorer social adjustments, and take higher doses of antipsychotics (Iasevoli, Balletta, Gilardi, Giordano, & De Bartolomeis, 2013).

6.5. The genetics of treatment resistance

Schizophrenia spectrum disorders are highly heritable; twin studies estimate that the heritability is 73% or more (Hilker et al., 2018). Joober et al. (2005) suggested that TR may also have a higher genetic loading than NTR, as they found that first- and second-degree relatives of patients with TR have a higher risk of developing schizophrenia spectrum disorders than the relatives of NTR patients. This has prompted a plethora of research into the genetic aetiology of TR.

At the time of writing, there have been at least twenty candidate gene association studies of TR: studies which have investigated polymorphisms in the brain derived neurotrophic factor (BDNF) gene (Anttila et al., 2005; Krebs et al., 2000; Zhang et al., 2013); the reelin gene (Goldberger et al., 2005); the g-protein signalling (RGS4) gene (Kampman et al., 2006); the disrupted-in-schizophrenia-1 protein (DISC 1) gene (Hotta et al., 2011); Cytochrome (CYP) P450 enzymes genes (van de Bilt et al., 2015); the dopamine system, including dopamine receptors and catechol-ortho-methyltransferase (COMT) enzyme, (Inada, Nakamura, & Iijima, 2003; Krebs et al., 1998; Ota et al., 2012; Teo et al., 2012); the serotonin system (Anttila et al.,

2007; Bilic, Jukic, Vilibic, Savic, & Bozina, 2014; Joober et al., 1999; Terzić, Kastelic, Dolžan, & Plesničar, 2015); and the immune system (Jia, Jayathilake, Zhao, & Meltzer, 2011; Lahdelma et al., 1998; Lin et al., 1998; Meged et al., 1999; Pinheiro et al., 2017). However, of the few identified candidate genes associated with TR, even less survived correction for multiple testing. In the region of the FAS gene, which is involved in inflammation, the rs7085850 marker was associated with TR (Jia et al., 2011). In the BDNF gene, 10 out of 34 polymorphisms were associated with TR (Zhang et al., 2013). Neither of these findings have been replicated. For more in-depth discussion of these candidate gene studies, see reviews by Elkis and Buckley (2016), Gillespie et al. (2017), and Pisanu and Squassina (2019).

It is now believed that individual polymorphisms have only minute effects on complex disorders, such as psychosis spectrum disorders, and therefore the sample sizes in most candidate gene studies are insufficiently powered to detect these effects. Thus, more attention has been given to genome-wide association studies (GWAS), which, by testing for the differences in single nucleotide polymorphism (SNP) frequencies between cases and controls, can identify risk or protective alleles from across the whole genome. GWAS with large sample sizes have been successful in identifying current or promising treatment targets in complex disorders; the Psychiatric Genomics Consortium (PGC) GWAS identified 108 independent genomic risk loci, 83 newly implicated in schizophrenia, when they used data from 36,989 schizophrenia patients and 113,075 healthy controls (Ripke et al., 2014). In the PGC's latest GWAS (an additional 11,260 patients and 24,542 controls) they found 145 independent loci, 50 of which were novel findings (Pardiñas et al., 2018).

To date, GWAS comparing TR and NTR patients have been small. Li and Meltzer (2014) had a sample of 174 patients (45% TR) and Koga et al. (2017) a sample of 84 patients (37% TR): no SNP met genome-wide significance in either GWAS. Instead, researchers hope to leverage the power of the PGC schizophrenia GWAS to investigate TR by summing the weighted effects of SNPs associated with schizophrenia and calculating a polygenic risk score (PRS or SZ-PRS). At one time, only polymorphisms which reached genome-wide significance were included

in a PRS, but it is now the convention to test multiple PRS, calculated using a range of P-value thresholds, and to use the threshold at which the PRS explains the highest proportion of variance. In testing the association between the SZ-PRS and TR findings have been mixed. Frank et al. (2015b) reported an association when using a P-value threshold of < 0.5 and when TR was defined using clozapine prescription. In Wimberley et al. (2017), there was no association regardless of the P-value threshold or definition of TR. Neither did Martin and Mowry (2016) find a significant association between the SZ-PRS and TR using a P-value threshold of < 0.5. Although this study is less comparable to the others, since they used the presence of psychotic symptoms and impaired functioning, after a continuous illness course and despite current treatment with antipsychotics, rather than two treatment failures, to define TR. In fact, the most recent study, by Legge et al. (2019), found an association between the SZ-PRS and TR using a P-value threshold of < 0.001. In all four studies, patients with TR had a higher SZ-PRS than NTR patients.

There is of course more to the genetic aetiology of TR than polymorphisms. SNPs are polymorphisms where the minor allele occurs in more than 5% of the population and are consequently referred to as common variants. Within the wider schizophrenia literature, there is evidence that risk variants for schizophrenia are not confined to common variants and are also found in rare chromosomal rearrangements including, *de novo* mutations, copy number variations (CNVs), rare single nucleotide variant (SNVs), and small insertion/deletions (indel) mutations (Rees, O'Donovan, & Owen, 2015). Research into rare variants associated with TR is still relatively rare. Taking a genome-wide approach, Martin and Mowry (2016) found that the total deletion burden (the total number of base pairs affected by copy number deletions) was higher in TR than NTR. But when Legge et al. (2019) tested candidate CNVs, included one previously associated with schizophrenia and one previously associated with intellectual disability, none were associated with TR. Using a gene set approach, Ruderfer et al. (2016) found an enrichment of rare mutations in gene targets for antipsychotics, and previously implicated pharmacogenetic genes, in TR patients when compared to NTR patients.

6.6. Summary and scope of thesis

In summary, studies to date have investigated a variety of clinical features in TR, including psychopathology, cognition, function, age of onset, comorbid diagnoses, and drug use. However, most of these studies have used cross-sectional designs and few have employed statistical methodology specifically designed to maximise prediction accuracy. In addition, there have been a range of studies investigating candidate genes, but none which have survived correction for multiple testing have been replicated. Two studies have shown that patients with TR have higher SZ-PRS, but since GWAS in TR have been small and underpowered, no PRS for TR (TR-PRS) has been able to be calculated. Finally, only a few studies have attempted to combine clinical and genetic data (Legge et al., 2019; Wimberley et al., 2017), and it likely that both forms of data are necessary to aid in the early identification of TR. There are substantial gaps in our understand of TR, which this thesis aims to address through the identification of clinical and genetic predictors of TR.

6.6.1. Aims

The aim of Chapter 8 was to conduct a systematic literature review in order to identify predictors of TR. This review only included studies which used a prospective follow-up design and focused on predictors that were identified at the first episode of psychosis, before antipsychotic treatment had commenced.

The aim of Chapter 9 was to build a model that could predict TR, using only clinical information collected prior to treatment with antipsychotics. The magnitude of effect between each predictor and TR was estimated using a logistic regression model. A prediction model of TR, containing only the most relevant predictors, was constructed using a penalised logistic regression.

The aim of Chapter 10 was to conduct a GWAS of TR that was adequately powered to detect the small effect sizes associated with SNPs. This estimated the effect of SNPs across the genome on TR, using two large case-control studies as discovery samples, and calculated a PRS for TR (TR-PRS). The TR-PRS was tested in two independent replication samples.

The aim of Chapter 11 was to take the findings from Chapters 8-10 and combine clinical and genetic data together in one analysis. Younger age of onset was the most consistent predictor of TR identified in my systematic review. In my own analysis, younger age of onset was one of only two predictors associated with TR, after controlling for all other variables, and it was selected for inclusion in the prediction model. In general, diseases with a high genetic burden tend to emerge in the early stages of a person's lifetime. Therefore, using path analysis I examined the relationship between age of onset and the TR-PRS.

Only clinical data and PRS are used in this thesis; neuropsychological data, candidate genes, and gene expression, may be important for predicting TR, but were not part of the analyses presented in this thesis.

METHODS

7. Method

7.1. Overview

The aim of this thesis is to investigate the clinical and genetic predictors of treatment resistance to non-clozapine antipsychotic medication, using a longitudinal cohort of patients who were recruited at their first episode of psychosis. The analytic sample used in this thesis is comprised of legacy data, aggregated as part of the STRATA-G Consortium.

7.1.1. STRATA

STRATA (Schizophrenia: Treatment Resistance and Therapeutic Advances) is an MRC-funded, UK-based consortium whose aim is to develop tests using information gathered (including clinical and demographic data, omics data, and MRI and PET imaging) to identify, in advance, which patients will fail to respond to non-clozapine antipsychotics. The rationale was that this approach may advance the progress of stratified medicine for schizophrenia and also afford insights into the mechanisms underlying treatment non-response. The Principal Investigator (PI) of STRATA was initially Professor Shitij Kapur at the Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London. The Chief Investigator (CI) was Dr James MacCabe who took over the role of PI in September 2016.

7.1.2. STRATA-G

Workstream 3 (WS3) of STRATA was entitled 'Predicting Response to Antipsychotic Medication using Clinical, Social and Genetic information'. The primary objective of WS3 was to collect DNA from pre-existing cohorts of first episode psychosis patients, to confirm pharmacogenetic predictors of response, in unbiased samples of patients who have been followed-up and assessed for response. WS3 was led by Professor James Walters and Professor Mick O'Donovan at the MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, and Professor James MacCabe and Professor Robin Murray at the IoPPN.

STRATA-Genetics (STRATA-G) is an international collaboration which grew out of its parent project, STRATA WS3. The STRATA-G collaboration brought

together data from well phenotyped prospective first episode psychosis cohorts with (1) DNA collected, (2) follow-up for a minimum of 1 year, and (3) sufficient data to determine treatment resistance/non-treatment resistance.

From October 2015, I corresponded, via email, teleconference, and Skype with Principal Investigators (PIs) about the inclusion of their data in STRATA-G. I co-wrote a Scientific Agreement. This agreement is intended as a memorandum of understanding between the scientists collaborating on STRATA-G, made in the spirit of co-operation and good faith, not as a formal or legally binding contract. It is based on the model of collaboration used successfully by the Psychiatric Genetics Consortium (https://www.med.unc.edu/pgc), and is based on the principle that scientists can best advance our shared goals through mutual collaboration. I coordinated and oversaw the implementation of material transfer and data transfer agreements to cover the formal matters pertaining to intellectual property and other legal issues. I was responsible for the transfer, storage, and processing of data for STRATA-G.

7.1.3. AESOP Follow-up

In two of the cohorts intended to be included in STRATA-G, the number of participants who had provided blood samples was low and so ethical permission was obtained to collect DNA from those who had previously participated in these studies. In Belfast, Dr Lina Homman followed up participants of the 'Northern Ireland First Episode Psychosis' and 'Resources for Genomics, Ireland' studies. In London, I followed up participants of the 'Aetiology and Ethnicity in Schizophrenia and Other Psychoses' (AESOP) study, and Ms Gemma Evans followed up AESOP participants who had originally been recruited in Nottingham.

I liaised with the research team and the PI, Professor Craig Morgan, to establish a list of potential AESOP participants. With section 251 approval (see below), this list was checked against hospital databases to update contact details, and appropriate GP/clinical team information. I also searched for instances of death and logged this if applicable. I also submitted an application to the Health and Social Care Information Centre (now known as NHS Digital) who provided updated contact details and instances of death for participants in this list.

As laid out in the study protocol I attempted to contact previous participants, firstly, by telephone using a script approved by the ethics committee, and secondly by letter to their postal address, asking for their help with this study and details about how to contact the research team if they wished to be involved. The participant could then either contact me by letter (I provided a stamped-addressed-envelope) or by telephone.

Once a participant had contacted me, we had discussed the research study further, and they had indicated that they wished to take part, the potential participants had two options: I could either visit their home or the participant could visit the research site to give a blood sample. The participant was offered £10 to compensate them for their participation. All travel expenses were reimbursed, or trains/taxis booked for the participant. I took informed consent from the participant, and then conducted a clinical interview, taking their medical history (which was confirmed by checking their clinical notes) and a blood sample.

7.2. Ethics

7.2.1. Ethical approval for STRATA WS3

The Research Ethics Committee (REC) at South Central - Oxford C approved STRATA WS3 and the follow-up of AESOP participants (reference: 15/SC/0021). A Health Research Authority (HRA) Confidentiality Advisory Group (CAG) also approved STRATA WS3 (reference: 14/CAG/1044).

STRATA and STRATA-G were conducted in compliance with the principles of the Declaration of Helsinki (1996), and in accordance with all applicable regulatory requirements including but not limited to the Research Governance Framework, Trust and Research Office policies and procedures, and those stipulated by the REC and CAG.

7.2.2. Ethical approval for AESOP Follow-up

For the AESOP follow-up, Section 251 approval was obtained to allow us to acquire updated contact information for these participants and to re-contact the participants, despite not having their explicit consent to do so. Section 251 refers

to section 251 of the National Health Service Act 2006 and its current Regulations, the Health Service (Control of Patient Information) Regulations 2002. As laid out by the HRA, "the NHS Act 2006 and the Regulations enable the common law duty of confidentiality to be temporarily lifted so that confidential patient information can be transferred to an applicant without the discloser being in breach of the common law duty of confidentiality. In practice, this means that the person responsible for the information (the data controller [AESOP PI and NHS Digital]) can, if they wish, disclose the information to the applicant [STRATA research team including myself] without being in breach of the common law duty of confidentiality. They must still comply with all other relevant legal obligations e.g. the Data Protection Act 1998. Approval also provides reassurance that that the person(s) receiving the information has undergone an independent review of their purposes and governance arrangements."

7.2.3. Ethical approval for STRATA-G

Ethical approval for STRATA-G was covered by the approvals for WS3. However, in addition the Scientific Agreement stated that each collaborator is to ensure that its work in STRATA-G complies fully with all applicable local, government and international laws, regulations and guidelines which are effective during the collaboration. They also needed to ensure that local ethical approvals are in place to cover the samples and data provided to the collaboration.

7.2.4. Material and data transfer agreements

Material and data transfer agreements (MTA and DTA, respectively) were put in place between King's College London (KCL) and Cardiff University (CU). A retrospective letter of agreement for the transfer of materials between the University of Sao Paulo and KCL was put in place. Nine three-way MTA/DTAs were put in place between KCL, CU, and the following institutions: Queen's University Belfast in the United Kingdom (UK); the University of Bologna in Italy; Istanbul University in Turkey; Lausanne University Hospital in Switzerland; the National Institute of Mental Health in the Czech Republic; Valdecilla Biomedical Research Institute at the Marqués de Valdecilla University Hospital in Santander, Spain; the University of Sao Paulo in Brazil; University College London (UCL) in the UK; Oslo University Hospital in Norway. For the Paris cohort, a declaration of

conformity with Reference Methodology MR03 was sent to the CNIL (National Commission on Informatics and Liberty) in accordance with French legislation.

7.3. Cohorts

7.3.1. King's College London (London, UK)

7.3.1.1. AESOP

The AESOP study (Aetiology and Ethnicity in Schizophrenia and Other Psychoses) is a multi-centre, naturalistic, prospective incidence and case-control study of first episode psychosis, conducted initially over a three-year period from September 1997 to August 2000. The study sample comprises: patients with an International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10; World Health Organization, 1992) diagnosis of F10-F29 or F30-F33, aged 16-65 years, who presented to secondary and tertiary services within tightly defined catchment areas in south-east London, Nottingham, and Bristol (Dazzan et al., 2005; Dean et al., 2018; Fearon et al., 2006; Kirkbride et al., 2006; Morgan et al., 2006; Zimbron et al., 2014). All participants, in centres in southeast London and Nottingham (UK), were invited to take part in a follow-up study, at approximately 10 years after baseline (Demjaha et al., 2017; Morgan et al., 2014; Revier et al., 2015) and those without bloods samples were invited to take part in a follow-up approximately 15 years after baseline for STRATA WS3. Treatment resistance/non-resistance was determined by Dr Arsime Demjaha (Demjaha et al., 2017) and myself.

7.3.1.2. GAP

The GAP study (Genetics and Psychosis) is a population-based incidence and case-control study of first episode psychosis, conducted initially over a three-year period from December 2005 to October 2008. The study sample comprises: patients with a ICD-10 diagnosis of F20-F29 or F30-F33, aged 18-65 years, who presented to secondary and tertiary services within a tightly defined catchment area in south-east London (Di Forti et al., 2015; Di Forti et al., 2009). Approximately 5 years after first contact for psychosis, a retrospective database search of clinical history was carried out using the Electronic Psychiatric Clinical Records (EPCRs) database (Ajnakina et al., 2017; Lally et al., 2016a). Treatment resistance/non-resistance was determined by Dr Olesya Ajnakina and Dr John Lally (Lally et al., 2016a).

7.3.2. Queen's University Belfast (Belfast, UK)

7.3.2.1. NIFEPS

The Belfast data is from two studies. The NIFEPS study (Northern Ireland First Episode Psychosis) is a naturalistic, prospective, incidence study, conducted initially over a two-year period from January 2003 and December 2004. The study sample comprises: patients with an Operational Criteria checklist for Psychotic Illness (OPCRIT; McGuffin, Farmer, & Harvey, 1991) diagnosis of first episode psychosis, aged 18–64 years, and living in Northern Ireland. All participants were invited to take part in follow-up studies, 1 year after baseline (Turkington et al., 2018) and approximately 13 years after baseline as part of STRATA WS3.

7.3.2.2. RGPI

The RGPI study (Resources for Genomics, Ireland) is a multi-centre, population-based, incidence study of first episode psychosis conducted from 2007. The study sample comprises: patients with a Diagnostic Statistical Manual version four (DSM-IV; American Psychiatric Association, 2000) diagnosis of schizophrenia, schizophreniform disorder, schizoaffective disorder or bipolar affective disorder with psychosis, aged 16+ years, who had Irish born grandparents, and who presented to psychiatric services in the region of the research centres (Casey & Corvin, 2008). All participants, recruited through psychiatric services in the region of Queen's University, Belfast, were invited to take part in a follow-up study, at approximately 9 years after baseline as part of STRATA WS3. Treatment resistance/non-resistance, in both samples, was determined by Dr Lina Homman.

7.3.3. University of Bologna (Bologna, Italy)

7.3.3.1. EUGEI

The Bologna data is from two studies. EUGEI (European Network of National Schizophrenia Networks Studying Gene-Environment Interactions) is a multi-centre, population-based incidence and case-sibling-control study of first episode psychosis conducted initially, in Bologna, over a four-year period from January 2011 to December 2014. The study sample comprises: patients with a ICD-10 diagnosis of F20-F33, aged 18-64 years, who presented to services within the catchment area (Jongsma et al., 2018). All participants, in Bologna, were invited to take part in a follow-up study, in 2016.

7.3.3.2. BoFEP

The BoFEP study (Bologna FEP) is an ongoing, naturalistic, prospective incidence study of first episode psychosis, conducted initially over an eight-year period from January 2002 and December 2009. The study sample comprises: patients with an ICD-10 diagnosis of F10–F29 or F30–F33, aged 18-64 years, who presented to services within the defined catchment area in West Bologna. All participants were invited to take part in a follow-up study, 1 year after baseline (Tarricone et al., 2012). Treatment resistance/non-resistance, in both samples, was determined by Dr Lina Homman.

7.3.4. Istanbul University *(*Istanbul, Turkey)

The Istanbul data is from an ongoing, hospital-based incidence study of first episode schizophrenia, conducted from 1996. This study is sometimes known as the First-Episode Schizophrenia Follow-Up Project and a proportion of this sample was included in EUGEI. The study sample comprises: patients with a DSM-IV diagnosis of schizophrenia, aged 15-45 years, who were experiencing an acute phase of their first psychotic episode and being treated as an inpatient (Üçok et al., 2016; Ucok, Polat, Cakir, & Genc, 2006; Ucok, Polat, Genc, Cakir, & Turan, 2004; Ucok, Serbest, & Kandemir, 2011). All participants were invited to take part in a follow-up study, 2+ year after their baseline (Üçok et al., 2016). Treatment resistance/non-resistance was determined by Dr Alp Üçok.

7.3.5. Lausanne University Hospital (Lausanne, Switzerland)

The TIPP study (Treatment and Early Intervention in Psychosis Program) is an ongoing, naturalistic prospective study of early onset psychosis, conducted from 2004. The study sample comprises: patients who meet threshold criteria for psychosis (defined by the 'Psychosis threshold' subscale of the Comprehensive Assessment of At Risk Mental States scale (CAARMS; Yung et al., 2002)), aged 18-35 years, who reside in the Lausanne catchment area (Alameda et al., 2017; Baumann et al., 2013; Golay et al., 2016). All participants enrolled in TIPP are invited to take part in follow-up studies, lasting 3 years after baseline. A subsample of TIPP patients were included in STRATA-G: those that participated either in a neurobiological research study developed by Prof Kim Do (Baumann et al., 2013),

and/or were part of PsyMetab or Psyclin studies (Choong et al., 2013; Choong, Solida, Lechaire, Conus, & Eap, 2008; Delacretaz et al., 2015; Quteineh et al., 2015; Vandenberghe et al., 2015). Treatment resistance/non-resistance was determined by Dr Romeo Restellini and Dr Luis Alameda.

7.3.6. University of Oslo (Oslo, Norway)

The TOP study (Thematic Organized Psychosis Research) is a naturalistic, prospective incidence and case-control study of first episode psychosis, conducted initially over a four-year period from May 2003 to July 2007. The study sample comprises: patients with a DSM-IV diagnosis of schizophrenia, schizophreniform disorder, schizoaffective disorder, psychosis not otherwise specified (NOS), delusional disorder, brief psychosis or major affective disorder with mood incongruent psychotic symptoms, aged 18-65 years, within 1 year of the start of their first adequate treatment with antipsychotic medication, who presented to outpatient and inpatient services within four University Hospitals in Oslo (Athanasiu et al., 2010; Faerden et al., 2008). All participants were invited to take part in a follow-up study, approximately 1 year after baseline (Faerden et al., 2013; Lange et al., 2014; Lyngstad et al., 2018). Treatment resistance/non-resistance was determined by Dr Carmen Simonsen and Professor Ingrid Melle.

7.3.7. French Institute of Health and Medical Research (Paris, France)

As part of the EUGEI, a multi-centre, population-based incidence and case-sibling-control study of first episode psychosis, subjects were assessed initially, in Créteil and Paris, over a two-year period from June 2012 to June 2014. The study sample comprises: patients with a ICD-10 diagnosis of F20-F33, aged 18-64 years, who presented to services within the catchment area (Jongsma et al., 2018). All participants, in Créteil and Paris, were invited to take part in a follow-up study, in 2017. Treatment resistance/non-resistance was determined by Dr Andrei Szöke and Jean-Romain Richard.

7.3.8. National Institute of Mental Health (Prague, Czech Republic)

The Early Stages of Schizophrenia study is a hospital-based incidence study of first episode schizophrenia, conducted initially over an unreported period of time. The study sample comprises: patients with a ICD-10 diagnosis of F20 or F23, aged 18-

35 years, who had less than 2 years of untreated psychosis, and who were hospitalised in a large general psychiatry hospital that serves Prague and part of Central Bohemia regions (Kolenic et al., 2018; Melicher et al., 2015; Mikolas et al., 2016; Spaniel et al., 2016). All participants were invited to take part in a follow-up study, 1 year after baseline. Treatment resistance/non-resistance was determined by Dr Lina Homman.

7.3.9. Marqués de Valdecilla University Hospital (Santander, Spain)

The PAFIP study (First Episode Psychosis Clinical Program) is an ongoing, naturalistic, prospective incidence study of first episode psychosis, conducted from February 2001. The study sample comprises: patients with an DSM-IV diagnosis of non-affective psychosis, aged 15+ years, who were referred from mental health services in the region of Cantabria. All participants were invited to take part in a follow-up studies, 3+ years after baseline (Ayesa-Arriola et al., 2018; Crespo-Facorro et al., 2007; Pelayo-Teran et al., 2008; Setien-Suero et al., 2018). Treatment resistance/non-resistance was determined by Noemí De La Fuente.

7.3.10. University of São Paulo (São Paulo, Brazil)

The Brazilian Wellcome Trust sample is a naturalistic, prospective incidence study of first episode psychosis, conducted between July 2002 and December 2004. The study sample comprises: patients with an DSM-IV diagnosis for psychotic disorder, aged 18-64 years, who had a first contact with mental health services due to a psychotic episode and who had been living in the defined geographical region of Sao Paulo for at least six months (Ayres et al., 2007; Martinho Jr et al., 2012; Menezes et al., 2007; Minatogawa-Chang et al., 2009; Schaufelberger et al., 2011). Participants who took part in the MRI part of the baseline study were invited to take part in a follow-up study, 2 years after baseline (Schaufelberger et al., 2011). Treatment resistance/non-resistance was determined by myself.

7.3.11. University College London (London, UK)

The West London Longitudinal First-Episode Psychosis Study a naturalistic, prospective incidence study of first episode psychosis, conducted from 1998 to 2008. The study sample comprises: patients with an DSM-IV diagnosis of psychosis, aged 16-50 years, who were presenting with a psychotic illness for the

first time and who had been receiving antipsychotic medication for less than 12 weeks. All participants were invited to take part in follow-up studies, 1, 3 and 5 years after baseline (Gutierrez-Galve et al., 2015; Gutierrez-Galve et al., 2010; Huddy et al., 2013; Huddy et al., 2007). Treatment resistance/non-resistance was determined by myself.

7.4. Defining Treatment Resistance

Treatment resistance (TR) was defined in two ways; participants could meet either of these two criteria to be classified as TR.

(I) Lifetime clozapine treatment was used as a simple marker of treatment resistance. This included patients who reported being prescribed or taking clozapine at a study visit or patients whose clinical notes stated that they were prescribed clozapine.

(II) The consensus definition developed by the *Treatment Response and Resistance in Psychosis (TRRIP) working group* (Howes et al., 2017).

The number of TR and non-treatment resistance (NTR) cases within each cohort are reported in Table 1. In two cohorts only lifetime clozapine treatment was available (Sao Paulo and UCL). As I am using legacy data there are instances when the criteria could not be applied as intended (e.g. in one cohort, their battery of tests did not include a measure of functioning). Table 2 shows how the TRRIP criteria were applied to each cohort. TR criteria had already been applied to AESOP and GAP so these definitions were retained as they were comparable to the TRRIP criteria (Demjaha et al., 2017; Lally et al., 2016a).

Our criteria deviated from the TRRIP criteria in three main ways; I did not apply their criteria concerning diagnosis, adherence, and symptom domain. Firstly, a confirmed diagnosis of schizophrenia was not included in the criteria for this study. Diagnosis at follow-up is not routinely collected in longitudinal studies and in my data 51% of follow-up diagnoses were missing. Only 25% of diagnoses at baseline were missing but in cohort studies, with only one study visit at baseline, diagnoses may not be valid. For example, DSM 5 criteria for schizophrenia require a

disturbance of six months and active symptoms for one month, so this diagnosis cannot be applied to participants who have been ill for less than six months at the time of the baseline assessment. Secondly, to be consistent across cohorts, patients did not need to be adherent to antipsychotics to meet the TRRIP definition of TR, as only two cohorts recorded adherence. In the Santander cohort, a subjective measure of adherence at one and three-year follow-ups was recorded (good vs. poor adherence). In the Lausanne cohort, antipsychotic blood levels were recorded. Thirdly, I did not specify which subclinical domain persistence symptoms must fall into. The TRRIP working group stipulate the need for subclinical specifiers (e.g. "positive," "negative," or "cognitive"). Not using subclinical specifiers means my findings are comparable to research published before the TRRIP guidelines.

Table 1. Number of participants stratified by treatment response and cohort.

Cohort	**NTR (%)**	**TR (%)**	**Total**
AESOP London	210 (73.43)	76 (26.57)	286
Belfast	138 (90.20)	15 (9.80)	153
Bologna	43 (86.00)	7 (14.00)	50
GAP London	216 (75.26)	71 (24.74)	287
Istanbul	33 (57.89)	24 (42.11)	57
Lausanne	250 (89.29)	30 (10.71)	280
Oslo	165 (97.62)	4 (2.38)	168
Paris	26 (81.25)	6 (18.75)	32
Prague	76 (54.29)	64 (45.71)	140
Santander	410 (86.32)	65 (13.68)	475
Sao Paulo	62 (95.39)	3 (4.62)	65
UCL London	364 (96.30)	14 (3.70)	378
Total	1992 (84.02)	379 (15.98)	2371

Table 2. Criteria used to define treatment resistance stratified by cohort.

Criteria/ Guidelines	Diagnoses at baseline recruitment	Number of antipsychotic medications	Type of antipsychotic medications	Duration of antipsychotic trial	Dosage during trial	Clinical response	Functional response	Adherence
AESOP (London, UK)	ICD-10 diagnosis of F10-F29 or F30-F33	2	NS	≥4 weeks	Daily dose of 400–600 mg of chlorpromazine equivalents	At least moderate severity on one or more positive symptoms as rated by SCAN (ICD-10)	NS	Recorded adherence to medication
NIFEPS & RGPI (Belfast, UK)	NIFEPS: OPCRIT diagnosis of first episode psychosis RGPI: DSM-IV diagnosis of schizophrenia, schizophreniform disorder, schizoaffective disorder or bipolar affective disorder with psychosis	2	At least two different antipsychotics	≥6 weeks	At least the mid-point of the licensed therapeutic range	At least moderate severity (a score of more than 70), as rated by the PANSS for at least 12 weeks	At least moderate functional impairment measured using the GAF scale (a score less than or equal to 51)	NS

EUGEI & BoFEP (Bologna, Italy)	EUGEI: ICD-10 diagnosis of F20-F33 BoFEP: ICD-10 diagnosis of F10–F29 or F30–F33	2	At least two different antipsychotics	≥6 weeks	At least the mid-point of the licensed therapeutic range	A rating of at least 70 on the PANSS	At least moderate functional impairment measured using the CGI scale	NS
Gap (London, UK)[1]	ICD-10 diagnosis of F20-F29 or F30-F33	2	NS	≥6 weeks	Daily dose of 400 mg of chlorpromazine equivalents	Little or no symptomatic improvement as determined by clinical case notes	NS	NS
Istanbul (Turkey)	DSM-IV diagnosis of schizophrenia	2	At least two different antipsychotics	≥6 weeks	At least the medium dosage of the therapeutic range	Persistent psychotic symptoms determined by a comparison of current and previous BPRS scores, alongside opinion of family member and treating psychiatrist	NS	NS

Lausanne (Switzerland)	Threshold criteria for psychosis as defined by the CAARMS scale	2	At least two different antipsychotics	≥6 weeks	At least the mid-point of the licensed therapeutic range	A rating of at least moderate severity as rated by the PANSS (minimum of 4 on at least 2 positive-subscale items or a minimum of 6 on 1 positive-subscale item) for at least 12 weeks	At least moderate functional impairment measured using the GAF scale	Excluded those with poor compliance

TOP (Oslo, Norway)	DSM-IV diagnosis of schizophrenia, schizophreniform disorder, schizoaffective disorder, psychosis not otherwise specified (NOS), delusional disorder, brief psychosis or major affective disorder with mood incongruent psychotic symptoms	2	At least two different antipsychotics	≥6 weeks	NS	A rating of at least moderate severity as rated by the PANSS (minimum of 4 on at least 2 positive-subscale items (or a minimum of 6 on 1 positive-subscale item) for at least 12 weeks	At least moderate functional impairment measured using the GAF scale	NS
Paris (France)	ICD-10 diagnosis of F20-F33	2	At least two different antipsychotics	≥6 weeks	At least the mid-point of the licensed therapeutic range	At least moderate severity, as rated by the PANSS, for at least 12 weeks	At least moderate functional impairment measured using the GAF scale	NS
Prague (Czech Republic)	ICD-10 diagnosis of F20 or F23	2	At least two different antipsychotics	≥6 weeks	At least the mid-point of the licensed therapeutic range	NS	NS	NS

Santander (Spain)	DSM-IV diagnosis of non-affective psychosis	2	At least two different antipsychotics	≥6 weeks	At least the mid-point of the licensed therapeutic range	A rating of at least moderate severity, as rated by the PANSS, for at least 12 weeks	At least moderate functional impairment measured using the GAF scale	NS
Sao Paulo (Brazil)[2]	DSM-IV diagnosis of psychosis	NS	NS	NS	NS	NS	NS	NS
UCL (London, UK)[2]	DSM-IV diagnosis of psychosis	NS	NS	NS	NS	NS	NS	NS

Abbreviations: NS, not specified; for all other abbreviations see List of Abbreviations on page 15.

NB: [1] Excluded those who were intolerant of antipsychotic medications or those who self-discontinued medication; [2] Only clozapine prescription could be used to determine TR.

7.5. Participants

Of the N=2371 participants, 37% had a schizophrenia diagnosis at baseline and 29% at follow-up. 67% have a confirmed psychosis-spectrum diagnosis at baseline and 46% at follow-up. In Table 3 I have reported how many participants had a diagnosis of schizophrenia, schizophrenia-spectrum disorders, and psychosis-spectrum disorders when data was non-missing.

Participants in the Oslo and Sao Paulo cohorts were excluded from the analysis in Chapter 9 as both cohorts has a low number of TR cases *and* a low proportion of TR cases (N=4 (2.38%) and N=3 (4.62%), respectively). Together, the small N and low proportion of cases meant that the software used for statistical analysis could not distinguish between TR and NTR.

Participants in the Belfast cohort who were genotyped on the Affymetrix chip were excluded from analysis in Chapter 10 and 11 due to a low overlap between the SNPs captured on this chip and SNPs captured on the Illumina chips used to genotype the other cohorts.

Participants were excluded from the analysis presented in Chapter 11 if they were missing data on age of onset or failed genetic quality control.

Table 3. Diagnoses of participants stratified by outcome.

	Baseline			Follow-up		
Diagnosis	NTR (%)	TR (%)	Total	NTR (%)	TR (%)	Total
Schizophrenia	714 (81.14)	166 (18.86)	880	572 (83.38)	114 (16.62)	686
Schizophrenia-spectrum	964 (83.32)	193 (16.68)	1157	751 (85.15)	131 (14.85)	882
Psychosis-spectrum	1,371 (85.69)	229 (14.31)	1600	946 (86.79)	144 (13.21)	1090

NB1: Schizophrenia-spectrum diagnoses include participants with a diagnosis of schizophrenia, schizoaffective disorder, or schizophreniform disorder. Psychosis-spectrum diagnoses include participants with a diagnosis of schizophrenia, schizoaffective disorder, schizophreniform disorder, delusional disorder, psychosis not specified as schizophrenia (NOS), acute and transient psychotic disorders, mania (with psychosis), or depression (with psychosis).
NB2: Follow-up refers to the last follow-up visit with non-missing data.

7.6. Phenotype data

7.6.1. Data pre-processing and cleaning

Collaborators transferred data files via the KCL and Cardiff secure file transfer systems. All original data files were stored by cohort. Copies were made and only copies were manipulated. Data preprocessing was conducted in STATA version 14 (Stata Statistical Software: Release 14, StataCorp LP., College Station, TX)

7.6.1.1. Pre-merge

Prior to merging the cohorts together, data was cleaned. This involved combining files which held different subsets of the data (e.g. baseline and follow-up), the removal of individuals with 100% missing data, the removal of variables with 100% missing data, and the creation of an unique identifier for STRATA-G, which was created by combining the original ID number and cohort name (e.g. EU163041ISTANBUL).

For each cohort, variables were renamed, labelled, formatted, and assigned value labels so that the coding of data was consistent across cohorts. Missing data flags (e.g. -99, -77, etc.) were edited to be consistent using the conventional missingness flags in STATA (e.g. “” for string variables and “.” for numerical variables). I also conducted basic quality check of variables; discussing with collaborators when impossible values occurred (e.g. a PANSS positive symptom scale score of 59 or a

birth date of 26/09/1862). When data inaccuracies could not be resolved, impossible values were changed to missing.

7.6.1.2. Merging data

Cleaned data files for each cohort were combined using the unique STRATA-G ID.

7.6.1.3. Post-merge

Additional variables were created for analyses (e.g. age of onset as discussed below). Again, a quality check of all variables was performed.

7.6.2. Predictors

Any variables recorded in more than one cohort, at baseline, were considered as potential predictors. I included both participant-related and illness-related variables. Details of how these variables were defined are below and the proportion of missing data is presented in Table 4 and the proportion of missingness stratified by cohort in Appendix A (Table 15 and Table 16). Descriptive statistics for each variable can also be found in the Appendix A (Table 17).

In addition, I calculated length of follow-up. Length of follow-up was recorded in years for GAP and Istanbul. However, for all other cohorts, length of follow was not recorded and, therefore, the difference between the baseline date and the furthest known follow-up date was calculated. This resulted in a length of follow-up measured in days, which was converted to years by dividing values by 365.25. For Bologna, follow-up dates were not recorded, but the original researchers assured me that all follow-ups occurred over a relatively short time period and that the 15 Nov 2016 could be used as a proxy follow-up date for all participants.

All the below variables were collected at baseline unless explicitly stated otherwise.

7.6.2.1. Participant-related variables

7.6.2.1.1. Accommodation

Accommodation status at baseline was recorded in Bologna, Lausanne, Oslo, and Paris, however, each used a different ordinal scale to capture this data. From this

information, I created a binary variable: in supported accommodation vs. independent living. I considered supported accommodation to include both practical support (e.g. assisted living, supervised living, institutions, care home) and financial support (e.g. council housing). I classified living with family as independent living; I did not have enough data to distinguish between participants living with family due to their illness or for other reasons (e.g. student or carer).

7.6.2.1.2. Age at baseline

Participant's age at the time of the baseline assessment was available for all samples.

7.6.2.1.3. Alcohol

Alcohol use at baseline was recorded in Belfast, Lausanne, Santander, and UCL, however each used a different scale. From this information, I created a binary variable: non-drinker vs. drinker. Alcohol use (yes vs. no) was recorded at baseline in Santander. DSM-IV criteria for alcohol use was recorded in Belfast (five categories: never used, abstention, use, abuse, and dependence). Never used and abstention were classified as 'non-drinker', while use, abuse, and dependence were classified as 'drinker'. The Case Manager Rating Scale (CMRS; Drake et al., 1990) for alcohol use was recorded in Lausanne (four categories: absent, light, moderate, and severe). Absent was classified as 'non-drinker', while light, moderate and severe were classified as 'drinker'. Alcohol use in the last six months, measured in number of units, was recorded in Oslo. Zero units in the last six months was classified as 'non-drinker', while more than zero units in the last six months was classified as 'drinker'. Alcohol use was also recorded in UCL (three categories: no, yes, and dependent). No was classified as 'non-drinker', while yes and dependent were classified as 'drinker'.

7.6.2.1.4. Body Mass Index (BMI)

BMI at baseline was recorded in Oslo and Santander. Weight in kilograms (kg) and height in centimetres (cm) at baseline was recorded in Belfast and GAP. The following formula was used to calculate BMI in Belfast and GAP: (weight/(height/100))/(weigh/100).

7.6.2.1.5. Cannabis

Cannabis use at baseline was recorded in Belfast, Bologna, GAP, Lausanne, Oslo, Santander, and UCL, however each used a different scale. From this information, I created a binary variable: cannabis vs. no cannabis. Cannabis use (yes vs. no) was recorded at baseline in Bologna, GAP, Santander, and UCL. DSM-IV criteria for cannabis use was recorded in Belfast (five categories: never used, abstention, use, abuse, and dependence). Never used and abstention were classified as 'no cannabis', while use, abuse, and dependence were classified as 'cannabis'. The CMRS for cannabis use was recorded in Lausanne (four categories: absent, light, moderate, and severe). Absent was classified as 'no cannabis', while light, moderate and severe were classified as 'cannabis'. Cannabis use in the last six months (yes vs. no) was recorded in Oslo. No use in the last six months was classified as 'no cannabis', while use in the last six months was classified as 'cannabis'.

7.6.2.1.6. Employment

Employment circumstances at baseline were recorded in Belfast, Bologna, GAP, Istanbul, Oslo, and Paris, however, each used a different ordinal scale. From this information, I created a binary variable: employed vs. unemployed. I considered 'unemployed' to include retired, economically inactive, rehabilitation welfare, disability benefit, and sick leave. Students were considered as 'employed' because in Istanbul and Oslo studying and working were grouped together.

7.6.2.1.7. Ethnicity

Ethnicity was recorded at baseline for all cohorts, apart from Belfast. Ethnicity in Belfast was recorded at the 10-year follow-up assessment. As this is a static trait, I used follow-up ethnicity in lieu of baseline ethnicity. For this analysis, I used three categories to define ethnicity: European, (e.g. White British, White other, White UK, White Irish, White Italian, White East European, Caucasian, European, American, and Gipsy), Black (e.g. Black Caribbean, Black African, Maghreb, African, Black), and Asian/Mixed/Other (e.g. Asian, Mixed Black, Mixed Other, Other, White and Black African, Arab, Filipino, Chinese, Bangladeshi, Indian, Middle East, Hispanic, and Latin American).

7.6.2.1.8. Gender

For all cohorts, gender was a binary variable and participants were either categorised as male or female.

7.6.2.1.9. Education qualifications

Highest educational qualification was recorded at baseline for AESOP, Bologna, GAP, Istanbul, Paris, and Prague. For this analysis, I used four categories to define highest educational qualification: None (e.g. no qualifications, primary school, school without qualification), Basic (e.g. GCSE, school with qualifications, O levels, junior/high school, secondary school), Further (e.g. first level on non-compulsory education, A levels, high school, Baccalaureate, vocation or college BTEC, NVQ, technical college), and Higher (university, undergraduate degree, postgraduate degree, professional).

7.6.2.1.10. Living

Living circumstances at baseline were recorded in AESOP, Belfast, Bologna, GAP, Lausanne, Oslo, and Santander, however, each used a different ordinal scale to capture this data. From this information, I created three binary variables: living with company vs living alone (e.g. spouse, children, family, parents, friends, other vs. alone), living with family vs. with non-family (e.g. partner, children, parents, other family vs. shared housing, friends), and living with parents vs. living with non-parents (e.g. parents vs. partner, children, other family, friends).

For living with company vs. living alone, a variable from Santander which grouped living alone without children and living alone with children together was categorised as living alone. For living with family vs. with non-family, the categories ‘alone’ and ‘other’ were not used. For living with parents vs. living with non-parents, data from Oslo was not used as they grouped parents and other family members into the same category.

7.6.2.1.11. Relationship

Relationship status at baseline were recorded in AESOP, Belfast, Bologna, GAP, Istanbul, Lausanne, Oslo, Paris, Prague, and Santander, however, each used a different ordinal scale to capture this data. From this information, I created two binary variables: current relationship and lifetime relationship; in a relationship vs

not in a relationship (e.g. married, steady relationship, cohabiting, civil partnership vs. single, separated, divorced, widowed, never married and not cohabiting) and ever been in a relationship vs. never been in a relationship (e.g. married, steady relationship, cohabiting, civil partnership, separated, divorced, widowed vs. single, never married and not cohabiting)

7.6.2.1.12. Tobacco

Tobacco use (yes vs. no) was recorded at baseline in Belfast, Bologna, GAP, Santander, and UCL. From this information, I created a binary variable: non-smoker vs. smoker.

7.6.2.1.13. Years in education

Number of years in education was recorded at baseline for Belfast, Bologna, GAP, Istanbul, Oslo, Paris, Prague, and UCL.

7.6.2.2. Disease-related variables

7.6.2.2.1. Age at onset

Age of onset was considered to be the participant's age when psychotic symptoms first occurred. Age of first psychotic symptoms was recorded at baseline for AESOP, Istanbul, Lausanne, Oslo, and Santander. If this variable was not available, the participant's age when they first presented to clinical services for psychosis was used. Age of first presentation to clinical services for psychosis was recorded at baseline for Belfast, GAP, and Paris. Age of first presentation to clinical services for psychosis was recorded at the five year follow-up for Bologna and since this is considered a static trait I included this data. As date of first presentation to clinical services is likely to be systematically later than date of first psychotic symptoms, I applied a correction to age at first presentation to clinical services when it was used instead of true age of onset. One cohort included in STRATA-G, AESOP, collected data on both age at first presentation and age of first symptoms. The mean difference between these ages was 0.547 years. Therefore, the estimated age of onset, in the absence of the variable, equalled age at first presentation to clinical services minus 0.547 years.

7.6.2.2.2. Brief Psychiatric Rating Scale (BPRS)

BPRS scores were recorded in Istanbul and Santander. The BPRS is designed to measure psychiatric symptoms, including hallucinations, depression, anxiety, and usual behaviour, across 24 items using a 1-7 Likert scale (1 equates 'not present' and 7 equates to 'extremely severe') (Overall & Gorham, 1962). The minimum score on the BPRS is 24 while the maximum is 168.

7.6.2.2.3. Duration of untreated psychosis (DUP)

DUP is the time in days between the first occurrence of psychotic symptoms and the start of antipsychotic treatment for psychosis. DUP was recorded in AESOP, Lausanne, Oslo, Paris, Santander, and UCL. DUP measured in weeks (Oslo, Paris, and UCL) was converted to days by multiplying by seven. DUP measured in months (UCL) was converted to days by multiplying by 30.417.

7.6.2.2.4. Family history of psychosis

Family history of psychosis (yes vs. no) was recorded in AESOP and UCL. Parental and family history of psychosis was recorded in AESOP, using the Family Interview for Genetic Studies (FIGS; Maxwell, 1992) and family history of schizophrenia was recorded in UCL, using the Diagnostic Interview for Psychosis (DIP; Castle et al., 2006). This included all known family members, and not just first-degree relatives.

7.6.2.2.5. Family history of mental health disorders

Family history of mental health disorders (yes vs. no) was recorded in AESOP and Belfast. Parental and family history of any mental health disorder was recorded in AESOP, using the FIGS. Only family history of any mental health disorder, using the FIGS, was recorded in Belfast. Within this variable I also included family history of psychosis, as described under 'Family history of psychosis'. Again, this included all known family members, and not just first-degree relatives.

7.6.2.2.6. Global Assessment of Functioning (GAF)

GAF scores were recorded in Belfast, GAP, Istanbul, Lausanne, Oslo, Paris, and Prague. The GAF is a scale, included in the DSM-IV, which is used to assess social, occupational, and psychological functioning (American Psychiatric Association, 2000). Individuals are given a score from 100 (extremely high functioning) to 1

(severely impaired). The GAF is often rated by focusing on symptoms (GAF-S) only or on functioning/disability (GAF-F) only; the GAF, as a single score, is the most serve of the GAF-S and GAF-F. GAF, as a single score, was recorded in Belfast, Istanbul, Lausanne, and Prague. In GAP, Oslo, and Paris, the GAF was recorded as two scores, and for each individual the most severe score was used. In Lausanne, the GAF was recorded two months after the baseline assessment, but I chose to treat it as the baseline score.

7.6.2.2.7. Positive and Negative Syndrome Scale (PANSS)

PANSS scores were recorded in Belfast, GAP, Lausanne, Oslo, and Prague. The PANSS is designed to measure symptoms of schizophrenia across 30 items using a 1-7 Likert scale for each domain (1 equates to 'absent' and 7 to 'extreme') (Kay, Fiszbein, & Opler, 1987). The minimum score on the PANSS is 30 while the maximum is 210. There are three subscales within the PANSS. The PANSS positive symptom subscale is used to rate positive symptoms of schizophrenia across seven domains e.g. delusions, hallucinations, etc. The scores for each item are summed so that the minimum score is 7 (all symptoms are absent) and the maximum score is 49 (all symptoms are present and extreme). The PANSS negative symptom subscale is used to rate negative symptoms of schizophrenia across seven domains e.g. blunted affect, stereotyped thinking, etc. The scores for each item are summed so that the minimum score is 7 and the maximum score is 49. The PANSS general psychopathy symptoms subscale is used to rate symptoms that are not covered by the positive or negative subscales. This subscale covers 16 domains e.g. somatic concerns, anxiety, depression, lack of judgement and insight, etc. The scores for each item are summed so that the minimum score is 16 and the maximum score is 112.

7.6.2.2.8. Scale for the Assessment of Negative Symptoms (SANS)

SANS scores were recorded in Belfast, Istanbul, Santander, and UCL. The SANS is used to rate negative symptoms in the following five domains: flat affect, alogia, apathy, anhedonia, and attention (Andreasen, 1983; Andreasen, 1989). Within each domain there are a number of individual items and one global item. Each item is rated on a 0-5 Likert scale (0 equates 'absent' and 5 equates to 'severe'). The SANS composite total is a sum of all SANS items apart from the global items (items: 1-

7, 9-12, 14-16, 18-21, and 23-24) the minimum score is 0 and the maximum score is 100. The SANS global summary score is a sum of all the five global items (items: 8, 13, 17, 22, 25), the minimum score is 0 and the maximum score is 25. SANS global summary scores were recorded in Belfast, Santander, and UCL. The SANS composite score was recorded in Istanbul. The SANS item 'inappropriate affect' is sometimes dropped from the composite total score because it does not correlate with the overall subscale score (Andreasen, 1982). This item was not recorded in both Belfast and UCL. As individual items were not available, I decided to use SANS global summary scores. Van Erp et al. (2014) was used to covert SANS composite total scores for participants from Istanbul into SANS global summary scores. The following equation was used: SANS global summary score = 1.0863+(0.2943*SANS composite total scores). The score for one individual was 25.8075, this was rounded down to 25.

7.6.2.2.9. Scale for the Assessment of Positive Symptoms (SAPS)

SAPS scores were recorded in Belfast, Istanbul, Santander, and UCL. The SAPS is used to rate positive symptoms in the following four domains: hallucinations, delusions, bizarre behaviour, and thought disorder (Andreasen, 1984). Within each domain there are a number of individual items and one global item. Each item is rated on a 0-5 Likert scale (0 equates 'absent' and 5 equates to 'severe'). The SAPS composite total is a sum of all SAPS items apart from the global items (items: 1-6, 8-19, 21-24, and 26-33). The SAPS global summary score is a sum of all the five global items (items: 7, 20, 25, 34). SANS composite total and global summary scores were recorded in Belfast, Istanbul, Santander, and UCL. In line with the SANS scores available, only SANS global summary scores were used.

Table 4. Potential predictors available at baseline.

Predictor	Missingness (%)	Model		Path Analysis
		Explanatory	Prediction	
Age of onset (years)	15.94	✓	✓	✓
Age at baseline (years)	7.80			
DUP (days)	28.38	✓	✓	
Female (vs. male)	15.94	✓	✓	✓
Body Mass Index	62.97		✓	
Being in a relationship (vs. single)	31.29		✓	
Ever being in a relationship (vs. never)	31.29			
Living alone (vs. with others)	50.91	✓	✓	
Living with non-parents (vs. with parents)	70.60	✓	✓	
Living with non-family (vs. with family)	66.81			
Living independent (vs. supported accommodation)	82.75			
Employed (vs. unemployed)	75.03			
Years in Education	69.17	✓	✓	
Cannabis (vs. no cannabis)	45.42	✓	✓	
Non-smoker (vs. smoker)	56.56	✓	✓	
Non-drinker (vs. drinker)	54.49	✓	✓	
PANSS Positive	71.49	✓	✓	
PANSS Negative	71.70	✓	✓	
PANSS General Psychopathology	71.87	✓	✓	
PANSS Total Score	72.33			
SAPS	60.14	✓	✓	
SANS	60.35	✓	✓	
GAF	74.91	✓	✓	
BPRS Total Score	77.77			
Family history of psychosis (vs. none)	83.80			
Family history of mental health disorders (vs. none)	82.12			
No qualifications (vs. reference)	72.54	✓	✓	
Basic education qualification (vs. reference)	72.54	✓	✓	
Further education qualification (reference)	72.54	✓	✓	
Higher education qualification (vs. reference)	72.54	✓	✓	
Non-European (reference)	32.22	✓	✓	
Black (vs. reference)	32.22	✓	✓	
Asian/Other (vs. reference)	32.22	✓	✓	

NB: Percentage missingness uses N=2371 as the denominator

7.7. Genotype Data

7.7.1. Genome-wide association study (GWAS): training data

Dr Antonio Pardiñas conducted a GWAS study of TR vs. NTR. The method he used is reported in detail in Chapter 10. In brief, two GWAS were conducted: the first compared patients with psychosis, who were not taking clozapine, to healthy controls and the second compared TR participants to healthy controls. In order to generate association statistics that reflect TR/NTR differences, a 'test for interaction' proposed by Altman and Bland (2003) was used. In other words, a 'virtual GWAS' was performed calculating the difference between odds ratios at each single nucleotide polymorphism (SNP). Summary statistics including z-scores and P-values were calculated.

7.7.2. Genome-wide association study (GWAS): test data

7.7.2.1. Genotypes

Collaborators transferred files containing genotypes for STRATA-G analysis via the KCL and Cardiff secure file transfer systems. DNA was sent to Cardiff via courier. Table 5 lists the genotyping arrays used to genotype the STRATA-G samples.

Table 5. Genotyping platforms used to genotype STRATA-G samples.

Cohort	Array Platform
AESOP London[1]	Infinium CoreExome-24 BeadChip (Illumina Inc., San Diego, CA, USA)
Belfast[1]	Genome-Wide Human SNP Array 6.0 (Affymetrix Inc, Santa Clara, CA, USA)
Bologna	Infinium HumanCoreExome-24 BeadChip (Illumina Inc., San Diego, CA, USA)
GAP London	Infinium HumanCoreExome-24 BeadChip (Illumina Inc., San Diego, CA, USA)
Istanbul[1]	Infinium HumanCoreExome-24 BeadChip (Illumina Inc., San Diego, CA, USA)
Lausanne	Infinium OmniExpress-24 v1.2 BeadChip (Illumina, San Diego, CA, USA)
Oslo	Infinium OmniExpress-12 v1_H BeadChip (Illumina, San Diego, CA, USA)
Paris[1]	Infinium HumanCoreExome-24 BeadChip (Illumina Inc., San Diego, CA)
Prague	Infinium OmniExpress-24 v1.2 BeadChip (Illumina Inc., San Diego, CA)
Santander[1]	Infinium OmniExpressExome-8 BeadChip (Illumina, San Diego, CA, USA)
Sao Paulo	Infinium OmniExpress-24 v1.2 BeadChip (Illumina, San Diego, CA, USA)
UCL London	Infinium OmniExpress-24 v1.2 BeadChip (Illumina, San Diego, CA, USA)

NB: [1] Additional samples genotyped for STRATA-G at Cardiff using the Illumina OmniExpress-24 v1.2 BeadChip (Illumina, San Diego, CA, USA)

7.7.2.2. Basic quality control (QC)

Basic QC was performed using Plink and R following the pipeline laid out by Anderson et al. (2010).

Each batch of called genotypes, within each cohort, were QC'd separately. Even though some Plink files contained individuals that were not going to be used in the STRATA-G analysis (e.g. healthy controls), these participants were retained during QC and imputation to improve accuracy of imputed genotypes.

First, sex was estimated using the homozygosity rate of the X chromosome (male sex is defined as $F \geq 0.8$ and female sex is defined as $F \leq 0.2$). When individuals have an ambiguous sex (e.g. > 0.2 and < 0.8) it is assumed that a genotyping error occurred, and these individuals were removed from further analysis. Individuals were also removed if their sex did not match their self-reported gender.

Variants in which the genotyping has performed poorly can be identified by assessing their call rates over the whole dataset (Turner et al., 2011). Single nucleotide polymorphisms (SNP) with a low call rate are more likely to contain genotyping errors, and thus I excluded SNP with < 95% call rate from further analyses.

Similarly, to the previous step, a large proportion of SNP assays failing on a single individual may be indicative of a poor-quality DNA sample, which leads to aberrant genotype calling. I removed individuals with < 95% call rate from further analysis.

To define an 'unrelated' subset of individuals, I estimated principal components that took relatedness into account. First, regions in long range linkage disequilibrium (LRLD) were removed (Price et al., 2008). I then used the R package PC-Air to estimate principal components (Conomos, Miller, & Thornton, 2015). PC-Air performs Principal Components Analysis on genome-wide SNP data for the detection of population structure in samples that may contain known or cryptic relatedness. Unlike standard PCA, PC-AiR accounts for relatedness in

the sample to provide accurate ancestry inference that is not confounded by family structure.

Using these principal components, I then estimating relatedness, taking ancestry into account, using the R package PC-Relate (Conomos, Reiner, Weir, & Thornton, 2016). PC-Relate uses the ancestry-representative principal components to adjust for population structure and accurately estimate measures of recent genetic relatedness, including the kinship coefficient. Kinship values were used to identify groups of individuals with different degrees of relatedness (Manichaikul et al., 2010). Identity-by-state (IBS) was calculated by examining, for each pair of individuals, the average proportion of shared alleles, excluding the sex chromosomes. When individuals share more alleles than expected by chance they are thought to be related and the proportion shared is proportional to relatedness. Individuals with IBS > 0.1 where flagged as related and individuals with IBS > 0.9 were assumed to be duplicates and removed from further analysis. For smaller batches of samples PC-Air/PC-relate could not be used and so the KING-robust method was used to create kinship estimates (Manichaikul et al., 2010).

7.7.2.3. Imputation & post-imputation QC

After I finished the basic QC pipeline, Dr Pardiñas uploaded genotypes to the Michigan Imputation Server (Das et al., 2016) where missing genotypes were imputed using the Haplotype Reference Consortium panel (McCarthy et al., 2016). Imputation infers the genotype at a particular SNP, which was not directly genotyped. When genotypes have been generated using different genotyping arrays, imputation is a method of standardising the data.

Dr Pardiñas then applied post-imputation filters (INFO > 0.8; Probability threshold > 0.9; Missingness < 5%; Hardy Weinberg Equilibrium P-value > 10^{-6}). Following imputation, the data contains a large number of variants and quality control is required again to remove missing, poor quality, or incorrectly assigned genotypes (Reed et al., 2015). The 'INFO' score is a measure of the quality of imputation (Coleman et al., 2016); usually a metric between 0 and 1, where 1 indicates a SNP has been imputed with a high degree of certainty. Here, we excluded any SNPs with an INFO score < 0.8. The probability threshold refers to probability assigned

to an imputed genotype. For a given SNP, there are three possible genotypes, and during imputation probabilities are assigned to each genotype. If the three probabilities, for the three genotypes, all fall below 0.9, we excluded that SNP. Variants with low certainty are marked as missing, and therefore not excluded when thresholding using the INFO score, so we further excluded all SNPs missing in more than 95% of participants. Finally, thresholds that identify missing variants do not necessarily remove miscalled variants, therefore we removed variants that deviated from Hardy–Weinberg equilibrium (HWE) (Coleman et al., 2016). HWE describes the relationship between allele frequencies and genotype frequencies. These are assumed to be constant over generations if there is no selection, mutation, or migration (Marees et al., 2018). By comparing the observed genotypes to the expected genotypes at each locus, using a χ^2 goodness-of-fit test, departures from HWE can be identified. Departures are assumed to indicate either the presences of evolutionary selection or genotyping errors. We excluded SNPs where the HWE test had a P-value < 10^{-6}.

7.7.2.4. Polygenic risk score (PRS)

A PRS is the sum total of weighted individual SNPs, which reach a certain significance threshold, and is a measure of genetic liability to clinical phenotypes. An additive model was used to sum SNPs, where each SNP is represented as the corresponding number of minor alleles (0, 1, or 2), and each SNP is weighted by its effect size. The log-odds of TR increases linearly as a function of the number of risk alleles. For Dr Pardiñas' work presented in Chapter 10, PRSice-2 was used to determine the significance threshold for the 'best-fit' PRS, rather than relying on the genome-wide significance threshold of P < 10^{-8} (Ripke et al., 2011) or including all SNPs (Dudbridge, 2013).

However, for the work presented in Chapter 11, an alternative method was used. It has been suggested that the power of polygenic risk scores increases with the number of included SNPs (Dudbridge, 2013), if shrinkage parameters are correctly applied. Therefore, rather than relying on the genome-wide significance threshold of P < 10^{-8} (Ripke et al., 2011), a PRS including all SNPs should be used. However, standard approaches to calculating PRS involve excluding markers in linkage disequilibrium (LD). Vilhjalmsson et al. (2015) suggest that, by excluding these

markers, the prediction accuracy of the PRS is less than the heritability explained by these SNPs. Therefore, PRS were calculated using LDpred-inf (Vilhjalmsson et al., 2015). LDpred-inf is a Bayesian method which estimates a posterior mean causal effect size from GWAS summary under a Gaussian infinitesimal prior. This is combined with LD information from a reference panel.

A SYSTEMATIC REVIEW OF PREDICTORS OF TREATMENT RESISTANCE

8. Predictors of Treatment Resistant Schizophrenia: A systematic review of prospective observational studies

8.1. Introduction

For approximately a third of patients with schizophrenia, standard antipsychotic medications do not adequately alleviate their psychotic symptoms (Conley and Kelly, 2001). This subgroup is termed treatment resistant schizophrenia (TRS). The most common clinical and research criteria used for TRS is the failure to respond to two trials of non-clozapine antipsychotics, of adequate dose and duration (Howes *et al.*, 2017, Suzuki *et al.*, 2011).

Patients with TRS have higher rates of unemployment, a worse quality of life, and poorer social and occupational functioning than people who respond to treatment (Iasevoli *et al.*, 2016). Researchers have estimated that the direct healthcare costs for TRS in the US is 3-11-fold higher than for the schizophrenia population as a whole, with multiple hospitalizations accounting for a large proportion of this cost (Kennedy *et al.*, 2014). In England, 25-50% of the National Health Service's (NHS) £11.8 billion mental health budget is allocated to schizophrenia services and TRS is thought to contribute a large proportion of these costs (Andrews *et al.*, 2012, Killaspy *et al.*, 2013).

Clozapine is the only antipsychotic recommended for TRS and is more effective than other antipsychotics in alleviating psychotic symptoms in patients with TRS (Kane *et al.*, 1988, Siskind *et al.*, 2016, Taylor, 2017). However, owing to its adverse effects, clozapine is only licenced in the UK (NICE, 2014) and most other developed countries (Warnez and Alessi-Severini, 2014) as a third-line treatment. Nevertheless, evidence suggests that TRS is often not recognised promptly, and that clozapine is offered after a delay of some years or not at all. According to treatment guidelines, the earliest that patients can be diagnosed with TRS, and prescribed clozapine, is 12 weeks after commencing antipsychotic treatment; however, Howes *et al.* (2012) report an average delay of 3.9 years, suggesting that there is considerable scope to shorten this period of inadequate treatment. Furthermore, patients with a shorter delay before clozapine initiation show a better symptomatic response to clozapine (Yoshimura *et al.*, 2017).

Thus, there is a need to identify patients - who are likely to develop TRS - earlier in the course of their illness and expedite their access to specialist treatment; this may require moving beyond the current definition of TRS towards criteria based upon predictors and biomarkers, which quantify a patient's risk of developing TRS. If predictors of TRS can be identified, they may be useful in three ways: firstly, to identify TRS patients earlier in treatment so that they can be offered effective treatments earlier; secondly, to identify patients for clinical trials of interventions for TRS; and thirdly, to improve our understanding of the aetiology of TRS.

We present a comprehensive systematic review of all prospective observational studies in schizophrenia populations, which report baseline predictors of TRS. We focused solely on prospective observational studies to draw clearer conclusions regarding the causal relationship between predictors and TRS in naturalistic settings over a long follow-up, and because only longitudinal studies can identify risk factors at first episode that might predict TRS.

8.2. Method

8.2.1. Inclusion/Exclusion Criteria

Studies were included if they met the following inclusion criteria: (1) participants were diagnosed with schizophrenia, schizophreniform disorder, schizoaffective disorder, and/or a psychotic disorder; we did not exclude studies that *also* included affective disorders or substance-induced psychosis, given the diagnostic uncertainly around the first episode of psychosis; (2) participants were followed from first episode or first treatment with antipsychotics; (3) the majority of participants were aged between 16 and 64 at baseline (we excluded studies that focussed exclusively on children or older adults); (4) data were collected prospectively from first episode; (5) the outcome was a categorical definition of TRS, established using longitudinal prospective medication history; and (6) a non-TRS comparison group was recruited and followed up in the same manner as the TRS group. Studies were excluded if (1) they were clinical trials, or if non-antipsychotic treatments, such as CBT or ECT, were administered as part of the study procedure; (2) the study focussed exclusively on early or late onset schizophrenia; or (3) inferential statistics measuring the association between

baseline variables and TRS were not reported, and our subsequent requests to the authors for unpublished data were unsuccessful.

8.2.2. Defining TRS

Only recently has attention been given to the standardisation of TRS criteria (Farooq *et al.*, 2013, Howes *et al.*, 2017, Lee *et al.*, 2015, Suzuki *et al.*, 2012); therefore, we did not restrict studies to one definition of TRS. We did, however, only include studies with a categorical definition of TRS to capture the key underlying concept - at least two treatment failures - and differentiate TRS from relative measures of response/nonresponse. If patients took clozapine at follow-up, we inferred that they met criteria for TRS. Clozapine prescription is likely to underestimate the true proportion of patients with TRS (Howes *et al.*, 2012), but it is a pragmatic criterion, since clozapine is only used for TRS, except in very rare indications (e.g. psychosis in the context of Parkinson's disease or for people who suffer severe side-effects to other antipsychotics).

8.2.3. Literature Search

Studies were identified by searching Pubmed, PsychINFO (up to October 2017), Medline (up to October 2017), Embase (up to October 2017), and OpenGrey on the 1st November 2017. In addition, we examined the first 20 pages of Google Scholar using terms 'predictor AND treatment resistant AND schizophrenia' on 3rd January 2018. No restrictions were placed on the publication date, but searches were restricted to the titles and abstracts of papers (and subject headings in Medline, Embase, and PsychINFO), studies published in English, and studies using human participants. Search terms for Pubmed were as follows: "((treatment resistant) OR (treatment resistance) OR (treatment refractory)) AND (schizophrenia) AND ((longitudinal) OR (prospective))". Search strategies for the other databases can be found in Appendix B (Table 18). We screened the title and abstracts of all identified studies and then performed full text screening of all potentially eligible studies. Potentially eligible studies were cross-referenced; additional relevant studies were identified by hand-searches of the references, and by screening papers which had previously cited these studies. Each additional paper was also hand searched until no new studies were identified. When full text articles were not available, the corresponding author was contacted. Author SES conducted

the initial screening, with APK independently screening the studies identified through database searches and all studies identified through cross-referencing.

8.2.4. Quality Assessment

We followed the PRISMA guidelines for reporting systematic reviews (Liberati *et al.*, 2009) (Appendix B Table 23).

Study quality was assessed using the Newcastle-Ottawa Scale (NOS) for cohort studies (http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp). Eight items measure the selection, comparability, and outcome of each study. These items were modified for this review, for example, follow-up needed to have been longer than one year to score on the item concerning adequate duration of follow-up (Appendix B Table 19). Authors SES and APK independently rated each study on the NOS (Appendix B Table 20), any differences in rating were discussed between authors and final ratings were a consensus.

When available, we report adjusted hazard (HR) or odds ratios (OR) with 95% confidence intervals (95%CI) in parentheses, for predictors measured at baseline.

8.3. Results

A total of 12 studies were identified for inclusion in this review. Study screening is depicted in Figure 1 and a summary of the number of participants recruited into each study is presented in Table 6. Database searches identified 545 records, 293 of which were duplicates and removed. 252 records were screened and 248 were excluded. The main reasons for exclusion were: the study did not follow participants from first episode or first treatment with antipsychotics (31%), participants recruited after TRS had been identified (29%), and an outcome other than TRS was reported (23%). The remaining 4 records were examined in more detail, as were the 8 records identified by cross-referencing. Only duplicates were identified through Google Scholar. Of the 12 studies, 11 were published in peer-reviewed academic journals. One study was unpublished (Chan *et al.*, 2014), however, after corresponding with the authors, a full report was identified on the funding body's website containing enough information to be included in this review (https://rfs1.fhb.gov.hk/app/fundedsearch/projectdetail.xhtml?id=1363).

Of the 12 included studies, 8 presented original data and 4 presented data on additional exposures within the same cohort as a previous study, or a subset thereof. Of these eight, three were population cohort studies. Both Sorensen *et al.* (2014) and Wimberley *et al.* (2016b) used Danish population registers: data was extracted from multiple national databases and linked using a unique personal identification number. Additional analyses of Wimberley *et al.* (2016b)'s data tested whether urbanicity (Wimberley *et al.*, 2016a), the polygenic risk score for schizophrenia (PRS-SZ; Wimberley *et al.*, 2017), functioning (Horsdal *et al.*, 2017b), and C-reactive protein levels (Horsdal *et al.*, 2017a) could predict TRS. The third population cohort came from South Korea (Kim *et al.*, 2017). The remaining five studies analysed longitudinal first episode psychosis patient cohorts (Chan *et al.*, 2014, Demjaha *et al.*, 2017, Lally *et al.*, 2016, Meltzer *et al.*, 1997, Üçok *et al.*, 2016).

In population registries, a proxy definition of first episode psychosis is required. In the Danish studies, the first International Classification of Diseases (ICD; World Health Organization, 1993) diagnosis of schizophrenia was used to define the baseline cohort. The South Korean study used ICD diagnosis of schizophrenia and first use of antipsychotics to define the baseline cohort. When using diagnoses, the first episode is likely to be later in the disease course, when compared to cohort studies. Additional study characteristics, including information about recruitment, diagnoses, and criteria for TRS and non-TRS can be found in Appendix B (Table 21). The variables measured, and tested as predictors of TRS, varied considerably across studies, therefore this information is summarised in Table 7. Appendix B (Table 22) contains the unadjusted and adjusted OR/HR, when these were reported.

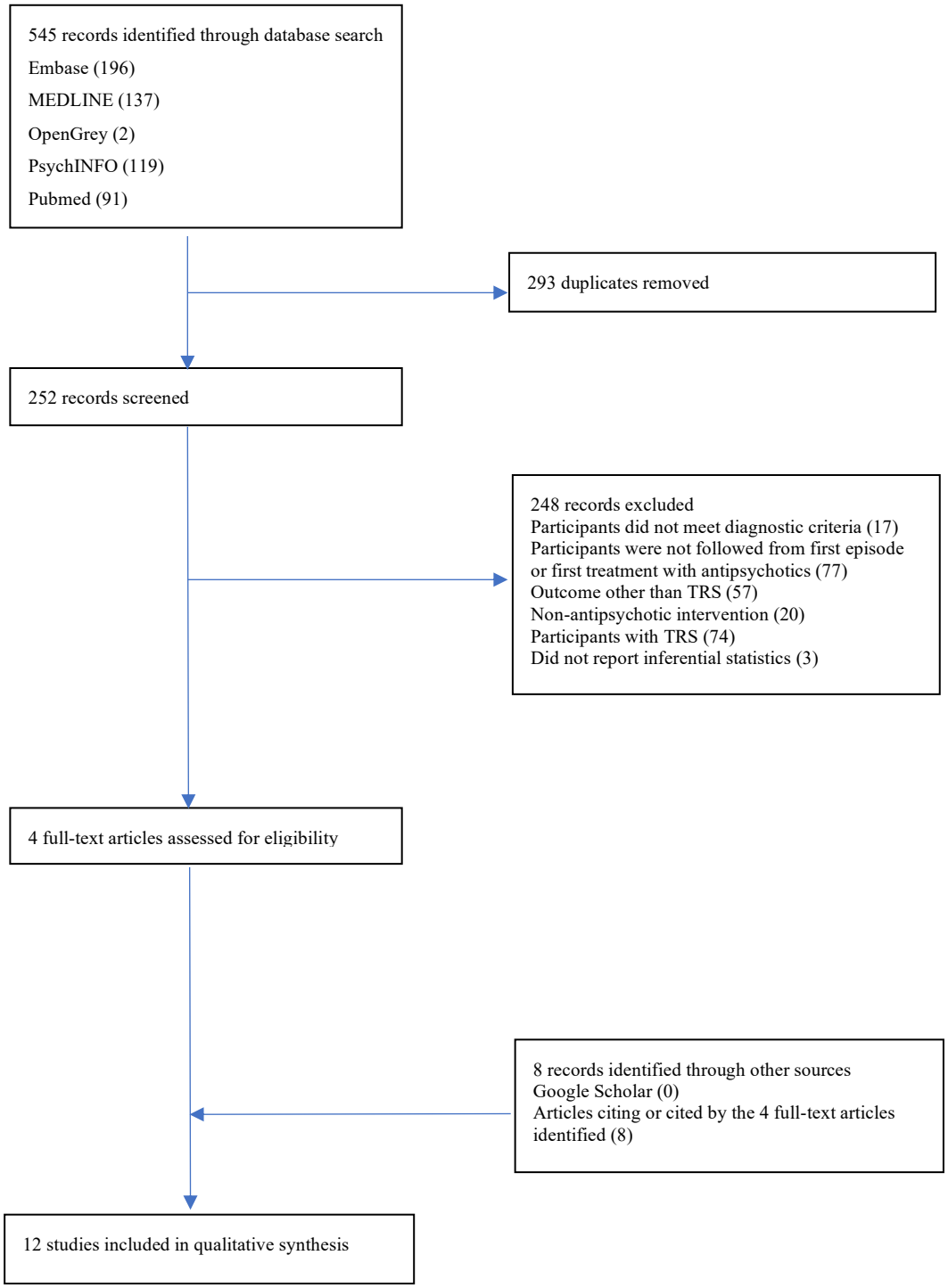


Figure 1. PRISMA Flow Diagram

Table 6. The twelve studies included in this review, with details on the number of participants recruited and the length of follow-up.

Study	Number of Participants							Length of Follow-Up (Years)
	Baseline	**Follow-Up**	**(%)**	**TRS**	**(%)**	**Non-TRS**	**(%)**	
Chan et al. (2014)	469	469	(100)	160	(34)	309	(66)	16
Demjaha et al. (2017)	557	274	(49)	62	(23)	212	(77)	10
Horsdal et al. (2017a)*	390	390	(100)	52	(13)	338	(87)	2
Horsdal et al. (2017b)*	3252	3252	(100)	359	(11)	2893	(89)	2
Kim et al. (2017)	114,749	NR		NR		NR		NR
Lally et al. (2016)	283	240	(85)	81	(34)	159	(66)	5
Meltzer et al. (1997)	322	322	(100)	196	(61)	126	(39)	4
Sorensen et al. (2014)	5968	5328	(89)	1223	(23)	4105	(77)	34
Üçok et al. (2016)	187	105	(56)	28	(27)	77	(73)	2
Wimberley et al. (2016a)*	13,349	13,349	(100)	2313	(17)	11,036	(83)	17
Wimberley et al. (2016b)	9332	8044	(86)	1703	(21)	6341	(79)	14
Wimberley et al. (2017)*	862	862	(100)	181	(21)	681	(79)	11

NB: NR, not reported; *, analysis of data also presented in Wimberley et al. (2016b)

8.3.1. Predictors of TRS

Chan *et al.* (2014) analysed a subsample of a first episode cohort who presented to mental health services over a five-year period and used clozapine prescription as a definition of TRS. As this was a case-control study including all patients with TRS and a ratio of 2 non-TRS patients for every TRS patient, the prevalence of TRS could not be calculated. The two groups were matched on baseline diagnosis. Chan *et al.* (2014) included age of onset, duration of untreated psychosis (DUP; days), duration of first episode, years of education, Premorbid Adjustment Scale (PAS) adult (19+ years) subscale score (Cannon-Spoor *et al.*, 1982), substance misuse history, and the number of relapses in the first three years, in a Cox proportional hazard regression. The model significantly predicted TRS (Chi-square = 66.11, *df* = 7, $p = < 0.0001$). While number of relapses in the first three years significantly predicted TRS, the only baseline predictors significantly associated with TRS were younger age of onset (HR = 0.88, 95%CI = 0.83-0.94) and poorer premorbid functioning (indicated by higher scores) according to the PAS (HR = 3.22, 95%CI = 1.43-7.23).

Demjaha *et al.* (2017) analysed data from the AESOP study, which recruited first episode patients over a three-year period and followed them up ten years later. The researchers entered gender, diagnosis, age of onset, negative symptoms, mode of onset, DUP (weeks), and ethnicity into a multivariate penalised logistic regression. The model selected five variables that predicted TRS: a diagnosis of schizophrenia at baseline (instead of psychotic depression; OR = 0.41, or psychotic mania; OR = 0.52), younger age of onset (years, OR = 0.97), higher severity of negative symptoms (OR = 1.09), an insidious mode of onset (instead of acute; OR = 1.28), and longer DUP (OR = 1.0013). Goodness-of-fit was measured using McFadden's pseudo R^2 and correct classification rates were measured using the Brier score. A McFadden's pseudo R^2 between .20 and .40 is considered a good model fit. The Brier score is used to evaluate predictive models; if the incidence of TRS is 23%, as estimated from Demjaha *et al.* (2017), a Brier score of 0 would be a perfect model while a score of .177 would be a non-informative model (Steyerberg *et al.*, 2010). Demjaha *et al.* (2017) reported a McFadden's pseudo R^2 of .10 and a Brier score of .146, suggesting that their model is not a good fit of the data nor is it a good classifier of TRS.

Kim *et al.* (2017), in their South Korean population cohort, estimated the cumulative incidence of clozapine use using the Kaplan–Meier method and log-rank test. They reported that younger age of onset predicted TRS. Unlike Chan *et al.* (2014) and Demjaha *et al.* (2017), Kim *et al.* (2017) examined age of onset categorically: defining younger age of onset as those aged between 15–20 years of age, and comparing them to a middle-onset group (21–44 years of age) and a late-onset group (45–64 years of age). Kim *et al.* (2017) also found, using the Walter–Elwood method (Walter and Elwood, 1975), a higher incidence of clozapine use, in those born during winter (December to February) when compared to those born in summer (June to August). This pattern remained true when stratifying season of birth by age of onset. Kim *et al.* (2017) reported no measures of overall model fit.

Lally *et al.* (2016) recruited first episode patients over a five-year period and used electronic medical records to follow them up five years later. They entered age of onset, Positive and Negative Symptom Scale (PANSS; Kay *et al.*, 1987) scores, Global Assessment of Functioning (GAF; Hall, 1995) disability score, and GAF symptom scores into a penalised logistic regression, controlling for living arrangements, employment status, and alcohol/substance misuse during the follow-up period. Lally *et al.* (2016) included the PANSS total score, the positive, negative and general psychopathology subscale scores, as well as two individual items: lack of insight and conceptual disorganisation. None of the PANSS or GAF variables predicted TRS. Age at first contact with mental health services was split into four categories: 18–20, 21–25, 26–30, >31 years. Only age of onset between 18 and 20 years, compared to all other age groups, significantly predicted TRS (OR = 2.49, 95%CI = 1.25-4.94). The authors did not report the overall model fit. Age of onset was subsequently stratified by gender and ethnicity. Age of onset, between 18 and 20, only predicted TRS in males (OR = 2.13, 95%CI = 1.35-7.23) or those of black ethnicity (OR = 3.71, 95%CI = 1.44–9.56).

Meltzer *et al.* (1997) recruited patients at first admission to hospital for schizophrenia or schizoaffective disorder and followed them up for approximately four years. The authors examined age of onset and gender in relation to TRS using

a two-way analysis of variance (ANOVA). Gender was not associated with TRS but younger age of onset was. As males had a younger age of onset than females, the researchers examined the associations between age of onset and gender in more depth using simple effects ANOVA. In the non-TRS group, males had a younger age of onset ($F = 6.6$, $df = 1$, $p < 0.01$), however, in the TRS group, there was no difference in age of onset between males and females. Meltzer *et al.* (1997) calculated the conditional probability of a patient having TRS given their age of onset. For those aged between 15 and 18 years old, the probability of developing TRS was between 32% and 38% for both males and females.

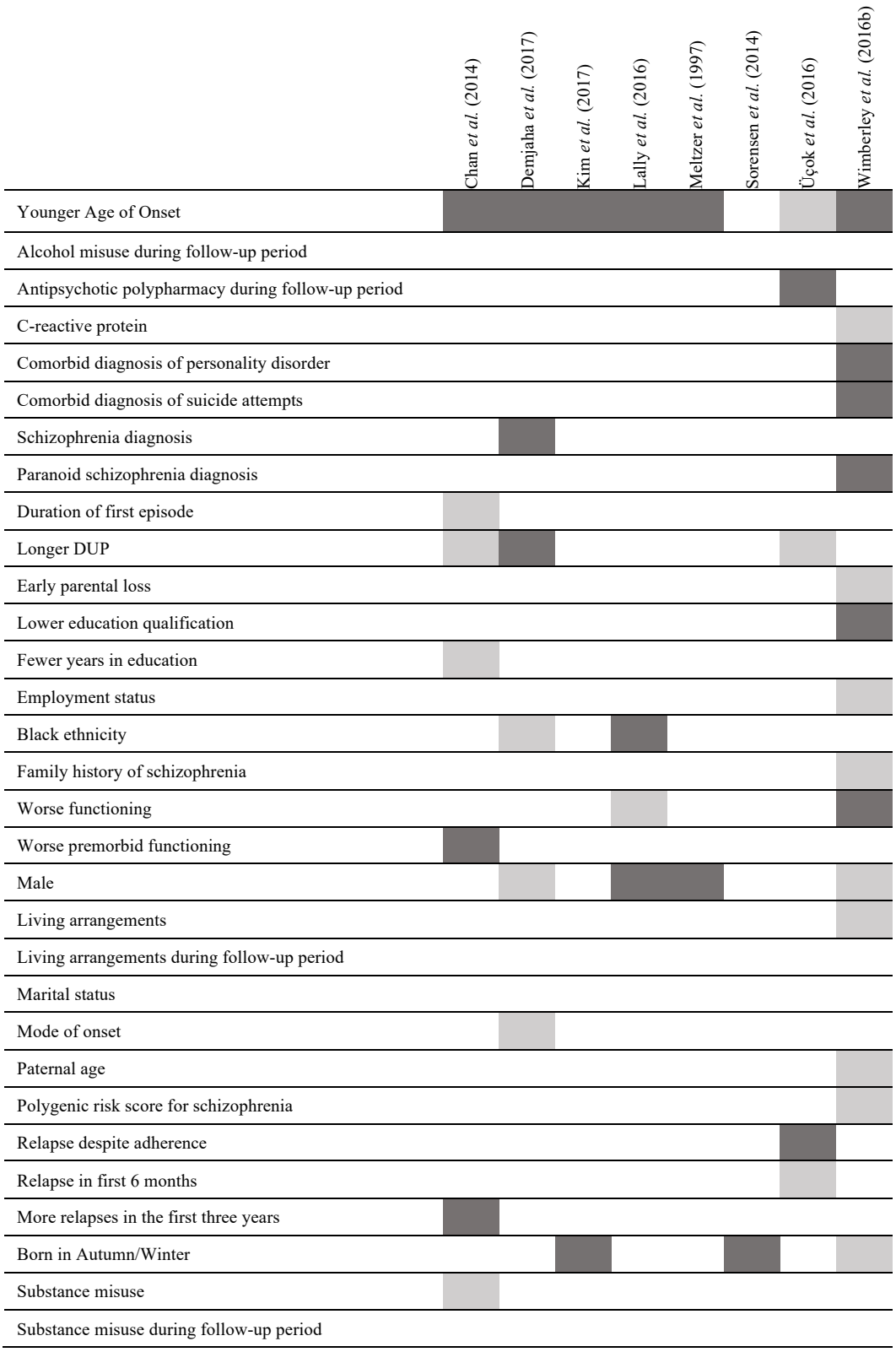
Sorensen et al. (2014), in their Danish population cohort, entered season of birth into a Cox proportion hazard regression adjusted for birth year and gender. The model did not significantly predict TRS. However, the authors found that being born in autumn (September to November), compared to spring (March to May), predicted TRS (HR = 1.24, 95%CI = 1.06-1.46). Unlike in Kim *et al.* (2017)'s study, being born in winter (December to February) failed to predict TRS.

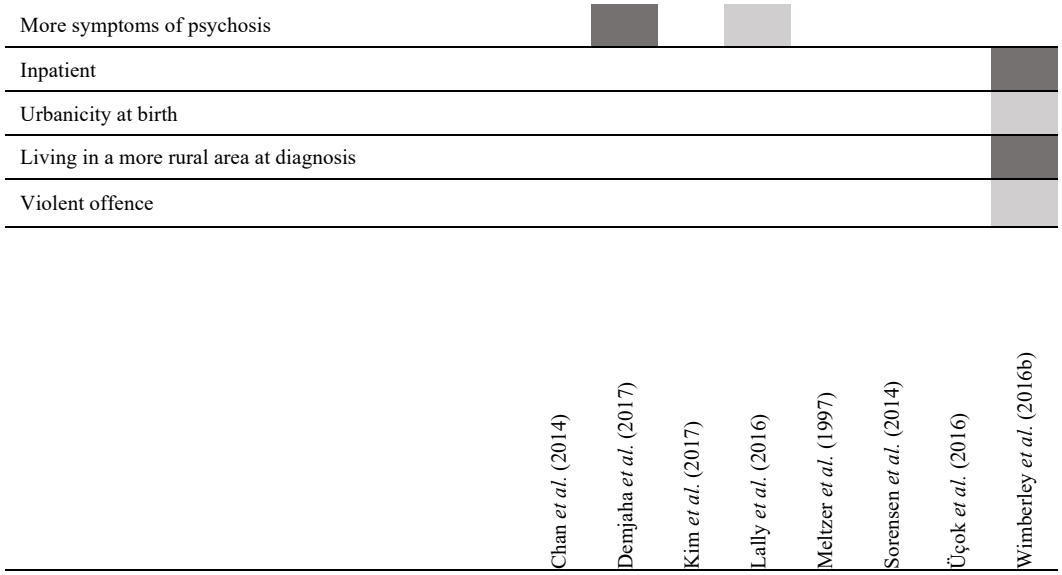
Üçok *et al.* (2016) analysed a subsample of patients recruited into an ongoing first episode schizophrenia study. Üçok *et al.* (2016) entered age of onset, DUP (days), first relapse despite adherence to antipsychotic treatment, relapse in the first six months, and antipsychotic polypharmacy during follow-up, into a logistic regression. The authors did not report the overall model fit. Only first relapse despite adherence to antipsychotic treatment and antipsychotic polypharmacy predicted TRS. No baseline variables predicted TRS.

Wimberley *et al.* (2016b), in their Danish population cohort, entered twenty-three variables into a Cox proportion hazard regression. These variables included: gender, age at first schizophrenia diagnosis as a proxy for age of onset, family history of schizophrenia in first-degree relatives, winter birth (December to March), paternal age, parental loss before the age of 18, living alone, conviction for a violent offence before first schizophrenia diagnosis, level of education, employment status, urbanicity at first schizophrenia diagnosis, admission to psychiatric hospital before first schizophrenia diagnosis, schizophrenia subtype (paranoid versus all others), comorbid psychiatric diagnosis before first

schizophrenia diagnosis, antipsychotic prescription in the year before first schizophrenia diagnosis, antidepressant prescription in the year before first schizophrenia diagnosis, and benzodiazepine prescription in the year before first schizophrenia diagnosis. Goodness-of-fit was measured using McFadden's pseudo R^2 and correct classification rates using Harrell's C statistic; a C statistic of 0.5 would be a non-informative model while a score of 1 would be a perfect model. Wimberley *et al.* (2016b) report a McFadden's pseudo R^2 of .027 and a Harrell's C statistic of .70, suggesting that this model is a good fit of the data and reasonable classifier of TRS. At baseline, younger age of onset (years, HR = 0.96, 95%CI = 0.95–0.97), living in less urban areas (rural versus capital area, HR = 1.44, 95%CI = 1.25–1.65), higher education (higher versus primary education, HR = 0.88, 95%CI = 0.79–0.98), psychiatric hospital admission at diagnosis (HR = 2.07, 95%CI = 1.87–2.29), having spent more than 30 bed-days in a psychiatric hospital in the year before diagnosis (HR = 1.54, 95%CI = 1.35–1.75), paranoid subtype diagnosis (HR = 1.24, 95%CI=1.13–1.37), comorbid personality disorder (HR = 1.24, 95%CI = 1.11–1.39), comorbid suicide attempt (HR = 1.21, 95%CI = 1.07–1.39), antipsychotic use (HR = 1.51, 95%CI = 1.35–1.69), antidepressant use (HR = 1.15, 95%CI = 1.03–1.29), and benzodiazepines use (HR = 1.22, 95%CI = 1.10–1.37), all predicted TRS. Data on additional exposures within the same cohort, or a subset thereof, were published separately. Lower levels of urbanicity (Wimberley *et al.*, 2016a) and severely impaired functioning (a GAF functioning score ≤ 30; Horsdal *et al.*, 2017b) predicted TRS, but the polygenic risk score for schizophrenia (PRS-SZ; Wimberley *et al.*, 2017) and C-reactive protein levels (Horsdal *et al.*, 2017a) did not predict TRS.

Table 7. The variables which have been tested as predictors of TRS in the twelve studies included in this review.





NB: grey squares, variables not significantly associated with TRS; dark grey squares, variables significantly associated with TRS; all analyses using the Wimberley *et al* Danish dataset were grouped under Wimberley *et al.* (2016b).

8.3.2. Subcategories of TRS

Some patients have little or no response to antipsychotic treatment from the onset of their illness, while others initially respond to medication and then later develop TRS. Two of the studies in our review reported comparisons between subgroups TRS patients; early-onset TRS was operationalised as meeting criteria for TRS from the onset of schizophrenia and delayed-onset TRS as meeting criteria after a period of symptomatic remission. Chan *et al.* (2014) found no differences, in demographics, clinical characteristics, or premorbid functioning, between early-onset TRS (N = 17, 11.64%) and delayed-onset TRS (N = 129, 88.36%). Lally *et al.* (2016) found no differences in demographics between the two groups, but the early-onset TRS group (N = 56, 70%) had a younger mean age of onset than the delayed-onset TRS group (N = 24, 30%).

8.4. Discussion

This review identified twelve research papers that examined predictors of TRS. Seven of the studies included in this review tested age of onset as a predictor, and six reported that younger age of onset predicted TRS. Given that multiple definitions of age of onset – age of onset of psychotic symptoms, age of first diagnosis of schizophrenia, age of first contact with mental health services – were reported and data was treated both continuously and categorically, this is a robust finding. Other potential risk factors, that have been identified by more than one study, include diagnosis, level of functioning, male gender, and season of birth.

A recent meta-analysis linked younger age of onset to multiple poor outcomes in schizophrenia: more hospitalizations, more negative symptoms, more relapses, poorer social/occupational functioning, and poorer global outcome (Immonen *et al.*, 2017). Many of these poor outcomes have also been associated with TRS. Immonen *et al.* (2017) found that males had a younger age of onset and, therefore, samples with a higher proportion of males tended to show stronger associations between age of onset and outcomes. In the studies included in this review, the association between age of onset and TRS is unlikely to be wholly confounded by gender, as the proportion of males ranged from 49% (Kim *et al.*, 2017) to 67% (Lally *et al.*, 2016) and the studies which controlled for gender still showed an

effect of age of onset (Demjaha *et al.*, 2017, Lally *et al.*, 2016, Meltzer *et al.*, 1997, Wimberley *et al.*, 2016b). In schizophrenia, age of onset has been thought to reflect genetic liability for the disease; younger age of onset has been associated with increased familial risk of schizophrenia (Byrne *et al.*, 2018, Hilker *et al.*, 2017). Could, therefore, TRS be the result of increased genetic risk? While Wimberley *et al.*, 2017 found no association between PRS-SZ and TRS, other work published by Frank *et al.* (2014) reports that an increased PRS-SZ is associated with TRS. In addition, rare copy number variations have been associated with both TRS (Martin and Mowry, 2015) and childhood-onset schizophrenia (Addington and Rapoport, 2009). Therefore, patients with TRS, who also have a younger age of onset, may have a more salient genetic influence than later onset cases, although further work is required to substantiate this claim.

This review complements previous reviews by Gillespie *et al.* (2017) and Carbon and Correll (2014). Gillespie *et al.* (2017) examined studies comparing patients with treatment-resistant to patients with treatment-responsive schizophrenia. They included all study methodologies, but excluded studies where treatment-responsiveness was defined solely as not meeting treatment-resistant criteria. Carbon and Correll (2014) examined studies identifying predictors of response and remission. The researchers focused on first episode psychosis studies where participants were followed up for five years. Some of the predictors of TRS, identified in this review, were found to be associated with less chance of response/remission by Carbon and Correll (2014), e.g. younger age of illness onset, poor premorbid adjustment, being male, lower level of education, living in a rural environment, diagnosis of schizophrenia, longer duration of untreated psychosis, poorer functioning, and worse psychopathology. However, Carbon and Correll (2014) also associated less chance of response/remission with being single, family history of psychosis, greater cognitive dysfunction, more family conflicts, and substance misuse; characteristics not identified as predictors of TRS. There was relatively little overlap between this review and Gillespie *et al.* (2017)'s review. In terms of studies included, only Meltzer *et al.* (1997)'s study was included in both reviews. In terms of characteristics associated with TRS, Gillespie *et al.* (2017) identified five neuroimaging studies, nine gene-association studies, and two studies of neurocognitive function, and these studies were not included in our review. The

examination of biological markers, associated with TRS, within longitudinal study designs is rare; this is understandable for genome-wide association studies, which require large sample sizes more easily acquired using a cross-sectional methodology. However, there is a clear gap in the literature investigating biological markers that change over time (for example, proinflammatory cytokines or differently methylated positions within the epigenome) and TRS as an outcome. In terms of neuroimaging research, a review by Nakajima *et al.* (2015) found only five studies which compared patients with TRS to non-TRS patients, none of which had identified neural correlates of TRS. McGuire and Dazzan (2017) highlight only one study where neuroimaging data predicted a six-year, non-remitting course of illness. Longitudinal imaging studies of TRS are still relatively rare and constitute another gap in the literature.

Of the studies included in this review, few identified characteristics of abnormal neurodevelopment as predictors of TRS, despite neurodevelopment changes being linked with schizophrenia. The neurodevelopmental theory of schizophrenia proposes that disrupted normal development, in utero or early infancy, leads to deficits in psychophysiological and neurological functioning in childhood or early adolescence, and eventually to prodromal or diagnostic symptoms of schizophrenia (Jablensky *et al.*, 2017, Murray *et al.*, 2017). Previous research has linked characteristics of abnormal development with TRS; higher rates of minor physical anomalies (Lin *et al.*, 2015), more neurological soft signs (de Bartolomeis *et al.*, 2018), poor verbal intelligence and fluency (Kravariti *et al.*, 2018), and poor verbal memory (de Bartolomeis *et al.*, 2013, Joober *et al.*, 2002). None of the studies in this review included variables measuring physiology during development or cognition at first episode. Only Chan *et al.* (2014) examined premorbid functioning, retrospectively using the PAS. They found no difference, between the TRS and non-TRS groups, in terms of functioning during childhood, early adolescence, or late adolescent. There was a difference in functioning after the age of 19 and subsequently, worse functioning predicted TRS in their final model. If educational attainment can be considered a proxy for development only lower level of education qualification was found to significantly predict TRS (Wimberley *et al.*, 2016b). Number of years in education was not predictive of TRS (Chan *et al.*,

2014). Abnormal neurodevelopment and neuropsychology have not been sufficiently investigated as potential predictors of TRS.

Our review has illuminated some gaps in the existing literature, where potential predictors have not been fully investigated, however, we believe our review has captured all published work and identified predictors that, with further study, may prove to be clinically useful in determining treatment for patients with schizophrenia.

8.4.1. Strengths and limitations

The main strength of this review is that we have focused solely on studies that included temporal forecasting (observations at baseline that are used to predict outcomes at follow-up), and as such eliminated recall bias and established a key component necessary for predictive models. All the studies included in this review are likely to be sufficiently powered to detect predictors of TRS. All the studies reported large sample sizes, and most followed participants for more than one year. Although no studies reported an *a priori* power analysis, and only Meltzer *et al.* (1997) reported an *ad hoc* power analysis, we believe lack of power is unlikely to explain these results.

When attrition reduces the sample size at follow-up of longitudinal studies, consequently, statistical power is also reduced. For the studies we have reviewed, that reported on participants lost to follow-up, it is unlikely that the low attrition rates introduced bias. In particular, many studies used Cox proportional hazard regression; an analytic method that not only takes into account that individuals lost to follow-up may develop TRS, but also that individuals may develop TRS after the study endpoint. However, TRS, in particular, may be biased by attrition. There is a case both that TRS patients may be more likely to drop out of research studies due to their higher severity of symptoms and worse social and occupational functioning, and that responders are more likely to drop out as they lose touch with clinical services, but we are not aware of any published studies examining attrition in relation to treatment response.

One limitation to consider, when discussing the findings from these studies, is that some patients may have been misclassified. None of the studies included in this review explicitly accounted for adherence to medication, therefore characteristics may be predicting nonadherence rather than treatment resistance. None of the studies measured antipsychotic plasma levels, therefore characteristics may be predicting sub-therapeutic drug plasma levels, as a consequence of nonadherence, noncompliance, or pharmacokinetics, rather than treatment resistance. McCutcheon *et al.* (2015) found that 44% of patients referred to an outpatient service for clozapine treatment had sub-therapeutic conventional-antipsychotic plasma levels. On the other hand, it is unlikely that TRS patients have been wrongly classified as responders because the long follow-up periods allow plenty of time for a diagnosis of TRS to be established. Most studies had follow-ups longer than four years; the average delay before being treated for TRS estimated by Howes *et al.* (2012). The definitions of TRS, used in these studies, are pragmatic criteria: any predictors identified by these naturalistic studies are generalisable to real-world, clinical settings where adherence, compliance, or drug plasma levels influence treatment.

The use of multiple definitions of TRS is a problem across all TRS literature; Suzuki *et al.* (2011) reviewed 33 studies of prospective studies of pharmacological interventions for TRS and found that all 33 definitions of TRS were different. Howes *et al.* (2017) reviewed 42 clinical trials and found only 2 studies which used identical criteria. In addition, some studies use clozapine prescription as a proxy for TRS. When clozapine is under-prescribed, supposed predictors of TRS may in fact represent predictors of clozapine initiation (e.g. clinicians' attitudes towards clozapine prescription). All of the studies, identified in this review, used existing data, not designed to examine TRS, and researchers had to established proxy definitions based on the data available to them. When evidence concerning predictors of TRS is not consistent, it can be hard to draw a clear conclusion about the validity of the predictor, yet when the evidence is consistent across studies, with different definitions, the predictor in question is highly likely to generalise to other cohorts and have clinical validity.

Finally, we must consider the statistical methodology used to establish predictors. A common misconception is that predictive accuracy can be inferred from explanatory accuracy. However, the two are different and should be assessed separately (Shmueli, 2010). Only three studies included in this review reported the overall model fit, and only two reported statistics that measure the predictive validity of the model. Additionally, in predictive modelling, variable selection and overfitting must be considered. Lally *et al.* (2016) and Demjaha *et al.* (2017) attempted to reduce overfitting by penalising regression coefficients. However, none of the studies used holdout data (training data), cross-validation, or external validation to evaluate the predictive power of models; the latter being the current 'gold-standard' approach. In terms of variable selection, the only methods reported were LASSO regression (Demjaha *et al.*, 2017) and step-wise selection using statistical significance (Chan *et al.*, 2014, Üçok *et al.*, 2016). Stepwise methods are no longer considered appropriate for explanatory models, but stepwise-type algorithms are very useful in predictive modelling (Shmueli, 2010), as long as the selection criteria rely on predictive power (e.g. Akaike Information Criterion) rather than explanatory power (e.g. statistical significance), as was the case in these studies. These methodological limitations must be taken into consideration when evaluating predictive models. The studies included in the review, on the whole, report analyses designed to identify explanatory variables of TRS. Future studies will need to use more robust prediction methods before moving from statistical prediction to clinical prediction.

8.4.2. Conclusion

The aim of this systemic literature review was to identify predictors of treatment resistant schizophrenia from prospective longitudinal studies. In choosing to focus exclusively on longitudinal studies, we have filled a gap in the existing literature, and hope that consolidating this information will be of use to researchers attempting to identify clinical predictors of TRS and the biological mechanisms causing TRS. We have identified earlier age of schizophrenia-onset as a robust predictor of TRS, with evidence that male gender, autumn/winter birth, poor premorbid functioning and rural upbringing may also contribute. We have also highlighted gaps in the literature namely, studies examining neuroimaging, immune, and genetic markers of TRS. Examination of biological markers,

particularly within the framework of a prospective longitudinal study, has the potential to go beyond simple prediction and add to our understanding of the underlying causes of TRS. In conclusion, while early identification of TRS is clinically important, we currently have very limited knowledge of its predictors.

A PREDICTION MODEL OF TREATMENT RESISTANCE

9. Clinical and Demographic Predictors Treatment Resistance: A prediction model

9.1. Introduction

Treatment resistance (TR) occurs in approximately 20-30% of patients with schizophrenia (Conley & Kelly, 2001). TR is most commonly defined as the continual existence of psychotic symptoms, despite two trials of antipsychotic medication, each lasting six weeks at a therapeutic dose (NICE, 2014). Clozapine is highly effective for TR (Agid et al., 2011), especially when prescribed earlier in the course of illness (Agid, Remington, Kapur, Arenovich, & Zipursky, 2007; Yoshimura et al., 2017), but research has shown a delay of around four years before it is initiated in clinical practice (Howes et al., 2012).

A prediction model capable of identifying TR earlier in the course of illness could help reduce this delay in clozapine initiation. Clinical prediction models investigate the relationship between future or unknown outcomes and baseline health states among people with a specific condition (Hemingway et al., 2013). They are used to calculate an individual's probability of developing a certain outcome. As a consequence, they can be used to inform clinical decision making (Lee, Bang, & Kim, 2016). Prediction models usually combine multiple parameters but must be simple and easy to use; there is often a trade-off between predictive power and usefulness in real-world settings (e.g. reducing the number of predictors).

It has previously been assumed that a statistical model, which explained a large proportion of the variance in outcome, would be good at predicting new cases, but this is often not the case. These so-called explanatory models are built to estimate unbiased causal associations between independent variables and an outcome, as accurately as possible within that specific sample. Usually the average relationship between two variables is presented together with a confidence interval and the result of a null hypothesis significance test (NHST). However, prediction of cases within the apparent sample does not guarantee good prediction of new cases in an external sample (Hastie, Tibshirani, & Friedman, 2009). Prediction models, on the other hand, are built with the primary aim of predicting new cases in future samples, but at the cost of lowing the interpretability of the model. These two different aims require different optimization criteria and thus different

methodologies (for a comprehensive review of the differences between explanatory and predictive modelling see Shmueli, 2010). Thus, machine learning and penalised maximum likelihood methods are often used for predictive modelling (Bzdok & Meyer-Lindenberg, 2018), of which penalised/regularized regression methods are well suited to predicting outcomes with a low prevalence, such as TR (Pavlou, Ambler, Seaman, De Iorio, & Omar, 2016). A further advantage of regularized regression methods is that, compared to more complex machine learning methods, the model coefficients and derived statistics can be interpreted more straightforwardly.

Jin, McCrone, and MacCabe (2019) examined the pharmacoeconomics of a prediction model for TR. They tested whether using a prediction model in clinical practice would be cost effective if individuals at-risk of developing TR were offered clozapine as second-line treatment. Taking into the account the cost of health and social care, the side effects associated with clozapine, and nonadherence, the authors concluded that a test with modest predictive power would be cost-effective. Data from prospective cohort studies is extremely valuable in prediction modelling, and the use of data collected prospectively at the first episode of psychosis provides a clinically practical baseline for predicting outcomes in psychosis. I recently conducted a systematic review of predictors of TR identified in prospective observational cohort studies (Smart, Kepinska, Murray, & MacCabe, 2019). Despite identifying twelve large prospective studies, only two reported statistics that measure predictive validity and all but one used statistical methods designed to identify explanatory variables of TR rather than predictive variables.

In the present study, my aim was to create two models: the first, an explanatory model to estimate the causal associations between first episode variables and TR; and the second, a predictive model that that could be used for early identification of TR and evaluated in terms of its predictive power to identify new cases of TR.

9.2. Method

9.2.1. Study Design and Participants

The 'Schizophrenia: Treatment Resistance and Therapeutic Advances' (STRATA) Consortium is a multi-disciplinary group of researchers working together to develop tests that can identify treatment resistance. As part of STRATA-Genetics (STRATA-G), legacy data was collected for combined analysis of observational, prospective, first episode psychosis cohorts. To be included in STRATA-G cohorts were required to have a minimum of one year of follow-up and genotypes (analysis of genetic data is reported elsewhere). Data included in this analysis comes from 10 cohorts in Czech Republic, France, Ireland, Italy, Spain, Switzerland, Turkey, and the United Kingdom (AESOP, GAP, UCL). Details of recruitment, follow-up, and where data has been previous published can be found in Chapter 7. Once cohorts had been combined, individual participants were excluded from this analysis if (1) missing data prevented the classification of participants as TR or non-treatment resistant (NTR), and (2) clozapine had been prescribed at the time of the baseline visit.

9.2.2. Treatment Resistance (TR)

I used three definitions of TR, depending on the data available in each cohort: either (1) prescription of or treatment with clozapine, or (2) treatment with two different antipsychotics at a therapeutic dose for a specified duration (e.g. each of at least 6 weeks' duration, with dosages in at least the mid-point of the licensed therapeutic range), or (3) persistent psychotic symptoms (e.g. as rated by the PANSS) and moderate functional impairment (e.g. as rated by the GAF) despite treatment with two different antipsychotics at a therapeutic dose for a specified duration of time. Participants not meeting any of these criteria were classified as NTR. A full list of the criteria used can be found in Chapter 7.

9.2.3. Predictors

Any variables recorded in more than one cohort, at baseline, were considered as potential predictors. I included both participant-related and illness-related variables (Table 4). Percentage of missing data, stratified by cohort, is in Appendix A (Table 15 and Table 16).

9.2.4. Other Variables

In addition, I collated length of follow-up and last known diagnosis recorded at a follow-up visit. When length of follow-up was missing, the difference between last known follow and baseline dates was calculated. Follow-up date was not available in the Bologna cohort, but as all follow-up visits occurred over a couple of months, the 15 Nov 2016 was used for all participants. When multiple diagnoses were recorded for follow-up visit, 'International Statistical Classification of Diseases and Related Health Problems' 10th Edition (ICD-10) diagnoses were used before 'Diagnostic and Statistical Manual of Mental Disorders' 4th Edition (DSM-IV) diagnoses, which were in turn used before self-reported diagnoses not attached to any specific diagnostic manual.

9.2.5. Statistical Analysis

I examined the correlation between predictors and the percentage missingness within each variable. When two continuous predictors had a Pearson's r greater than 0.8 or two categorical predictors had a Spearman's rho greater than 0.8, one was removed from all subsequent analyses, based on the authors knowledge of the literature. Data was missing systematically across study sites. I removed all predictors where more than 75% of the data was missing.

Data preprocessing was conducted in STATA version 14 (Stata Statistical Software: Release 14, StataCorp LP., College Station, TX) and analysis performed in R version 3.6.0 (R Core Team, 2009). Scripts used for analysis can be found on the Open Science Framework: https://osf.io/5sq3z/.

9.2.5.1. Explanatory Model (Logistic Regression)

For variables with less than 75% missingness, I imputed missing data. I used multiple imputation by chained equations (MICE), using the 'mice' package in R (van Buuren & Groothuis-Oudshoorn, 2011), and included the indicators for study site in the imputation model, without clustering, rather than multilevel MICE with a random intercept model. Multilevel MICE could not be applied to my data due to the low number of TR cases and the uneven distribution of observed data across cohorts. I compared MICE and multilevel MICE on a subsample of the data (consisting of a subsample of cohorts and variables) that fit the distributional

assumptions and found no difference in log odds and standard error (data not shown, available upon request). I ran MICE using 100 imputations, with 10 iterations in the burn-in period. Performance of MICE was assessed by plotting the residuals and convergence of chains, exploring the range of imputed values, and comparing observed and imputed data, following the guidance in Nguyen, Carlin, and Lee (2017). Each of the imputed datasets were analysed separately and then the results were pooled together based on Rubin's rules (Azur, Stuart, Frangakis, & Leaf, 2011; Rubin, 1987; Schafer & Graham, 2002). Regression coefficients were pooled by taking the average coefficient from all the imputed datasets and standard errors by combining the within imputation variance and the between imputation variance. Imputation model checking was performed (Nguyen et al., 2017).

I conducted a logistic regression to examine the associations between predictors and TR. Variable selection was based on a systematic literature performed by myself (Smart et al., 2019). This review included longitudinal studies which had examined the effect of variables collected at first episode on TR. Length of follow-up and cohort were included to control for differences between cohorts. I report the pooled estimates from univariable analyses and the multivariable model.

9.2.5.2. Prediction Model (LASSO Logistic Regression)

I conducted a 'Least Absolute Shrinkage and regression Operator' (LASSO) logistic regression to identify predictors of TR. LASSO is a form of regularised regression which can avoid overfitting when there is a large number of predictors. I preferred lasso over other regularization methods (ridge or elastic net) because it will select a sparse and, therefore, more clinically practical model. The model coefficients are shrunk towards zero, with some coefficients being assigned exactly zero to perform variable selection. I imputed missing data using the random forests imputation method with the 'missForest' package in R (Stekhoven & Buhlmann, 2012). Lasso regression counterbalances the known bias, that occurs due to overfitting (optimism), during random forest imputation (Lu & Petkova, 2014). Previous work by Dr Agbedjro found that random forests imputation outperformed mice when paired with a LASSO regression (Agbedjro, 2018). Random forests imputation accuracy was assessed using the out-of-bag imputation normalized root

mean squared error (NRMSE) for continuous imputed data, and the proportion of falsely classified entries (PFC) for the categorical imputed data set (in both cases, a value close to 0 indicates good performance and a value close to 1 poor performance; I considered a value around 0.5 to indicate a moderate performance). The tuning parameter lambda was estimated on the imputed data using a grid of 100 tuning parameters (λ) and by maximising the area under the receiver operating characteristic curve (AUC), using the 'caret' package in R. The model was obtained using repeated cross-fold validation (5 folds, 50 repeats). The one standard error (1 SE) rule (delivering an AUC within 1 SE of the maximum) was reported as simulations suggest that this model provides the best compromise between reliable variable selection and good prediction accuracy (Hastie et al., 2009). The model's apparent discriminative performance was evaluated using the AUC. When the AUC is 0.5, the model is no better than chance at predicting TR. I report the model sensitivity (number of true TR cases), specificity (number of true NTR), positive predictive value (PPV), negative predictive value (PNV) at the best risk threshold. A best threshold should be determined by the cost and benefits of an intervention (Steyerberg, 2008). As I did not have access to this information, I used the threshold that maximises the sum of sensitivity and specificity.

I conducted four sensitivity analyses. First, I repeated the model using only clozapine prescription to define TR. Second, I repeated the model using only participants with a diagnosis of schizophrenia (at the last known follow-up visit). Third, I repeated the model while oversampling TR cases. In prediction models of a binary outcome, a model that classifies all data into the majority class (NTR) can have good prediction accuracy. To prevent this, random oversampling uses all data, then randomly selects, with replacement, data from the minority class (in my case TR) until the two classes are balanced (Dubey, Zhou, Wang, Thompson, & Ye, 2014). Finally, I repeated the model using repeated 5-fold validation (5 folds, 50 repeats) to examine the discrimination performance measures within each cohort separately.

I internally validated the model using repeated cross-fold validation (5 folds, 50 repeats) to compute optimism-corrected performances (Harrell, Lee, & Mark,

1996)[1]. The difference between the performance estimates of the model developed on each training subsample and the test performance measures of these models applied to the left out of fold of data were averaged to obtain an estimate of the optimism. The optimism is then subtracted from the apparent performance measures of the model developed on the whole dataset to obtain the corrected measures. I present the corrected performance measures and recalibrated coefficients.

9.3. Results

From ten first episode cohorts, I collated data for N=2138 participants, N=372 (17.40%) were classified as TR and N=1766 (82.60%) as NTR (Table 8). Descriptive statistics for all independent variables are presented in Appendix A (Table 17).

I excluded five variables where more than 75% of the data was missing: family history of psychosis (83.80%), accommodation status (82.75%), family history of mental health disorders (82.12%), BPRS total score (77.77%), and employment status (75.03%). Of the remaining variables, age at baseline and age at onset were correlated (r = 0.96), as were PANSS total score and PANSS general psychopathy subscale score (r = 0.91), as well as current relationship and lifetime relationship (rho = 0.95). Lifetime relationship was also correlated with living with parents (rho = -0.82). I therefore removed age at baseline, PANSS total score, and lifetime relationship from all analyses. I also excluded living with family due to low cell count.

[1] Dr Agbedjro and I would like to thank Professor Frank E Harrell Jr for his advice on using repeated cross-validation to correct for optimism, see: https://discourse.datamethods.org/t/optimism-correction-for-lasso-cox/1297/4.

Table 8. Number of participants stratified by treatment response and cohort.

Cohort	NTR (%)	TR (%)	Total
AESOP London	210 (73.43)	76 (26.57)	286
Belfast	138 (90.20)	15 (9.80)	153
Bologna	43 (86.00)	7 (14.00)	50
GAP London	216 (75.26)	71 (24.74)	287
Istanbul	33 (57.89)	24 (42.11)	57
Lausanne	250 (89.29)	30 (10.71)	280
Paris	26 (81.25)	6 (18.75)	32
Prague	76 (54.29)	64 (45.71)	140
Santander	410 (86.32)	65 (13.68)	475
UCL London	364 (96.30)	14 (3.70)	378
Total	1766 (82.60)	372 (17.40)	2138

9.3.1. Explanatory Model

As BMI and relationship status at baseline had not been previously associated with TR, I did not include these variables in my explanatory model.

MICE performed well; plots of the residuals and convergence were acceptable. All imputed values fell within either a pre-determined range (e.g. the minimum score on PANSS positive symptoms subscale is 7, while the maximum score is 49, therefore all imputed scores must fall between 7 and 49) or within the range of nonmissing values (e.g. the lowest observed age at onset was 6, while the oldest observed age at onset was 63, therefore all imputed scores must fall between 6 and 63). The density of observed data and the imputed data was similar, as would be expected under the missing-at-random assumption.

In univariable logistic regressions, TR was associated with younger age at onset, male gender, fewer years in education, higher PANSS negative subscale score, higher PANSS general psychopathology subscale score, lower GAF score (worse functioning), black ethnicity (compared to European ethnicity), and possessing a further education qualification as the highest educational achievement (compared to possessing a higher education qualification).

In the multivariable logistic regression, when coefficients are adjusted for all other variables as well as cohorts and length of follow, TR was associated with younger age of onset (OR = 0.95; 95%CI: 0.91, 0.99; p = 0.005) and fewer years in

education (OR = 0.86; 95%CI: 0.70, 1.02; p = 0.048). Coefficients for both univariable and multivariable models are shown in Table 9.

9.3.2. Prediction Model

Missing data imputation performed moderately: the normalized root mean squared error (NRMSE) for the continuous variables was 52.32% and the proportion falsely classified (PFC) for the categorical variables was 12.28%.

In the 1 SE LASSO regression model, 9 of the 24 variables were selected. Younger age of onset, male gender, higher BMI, being single, living with non-parents, fewer years in education, drinking alcohol, lower GAF score, and achieving a higher education qualification, were associated with a higher probability of being TR (Table 10). This model predicted TR with an apparent AUC of 0.65 (McNemar's Test P-value < 0.001; see Figure 1) and a corrected AUC of 0.60 (Table 11 and Figure 2). The equation for the corrected model and a nomogram plot to work out the risk a new patient will develop TR can be found in Appendix C (Equation 1 and Figure 10, respectively).

The apparent AUC was 0.67 for the clozapine model, 0.66 for the schizophrenia model, and 0.67 for the oversampled model (Appendix C Table 24 and Table 25). Apparent AUC ranged from 0.60 to 0.76 across the ten cohorts (Appendix C Table 26). The corrected AUC was 0.62 for the clozapine model, 0.56 for the schizophrenia model, and 0.64 for the oversampled model (Appendix C Table 27 and Table 28).

Table 9. Results of multilevel logistic regressions.

	Univariable (unadjusted) models							Multivariable (adjusted[1]) model						
	Coefficients	**SE**	**95%CI**		**Odds Ratio**	**P-value**		**Coefficients**	**SE**	**95%CI**		**Odds Ratio**	**P-value**	
Intercept	NA							-1.74	2.26	-6.17	2.69			
Age of onset (years)	-0.04	0.01	-0.06	-0.08	0.96	< 0.001	***	-0.05	0.02	-0.09	-0.01	0.95	0.005	**
DUP (days)	0.00	0.00	0.00	0.00	1.00	0.36		0.0001	0.00	0.00	0.00	1.00	0.53	
Female (vs. male)	-0.35	0.14	-0.62	-0.87	0.70	0.01	*	-0.10	0.22	-0.53	0.33	0.90	0.66	
Living alone (vs. with others)	0.06	0.19	-0.31	0.67	1.06	0.77		-0.01	0.30	-0.60	0.58	0.99	0.96	
Living with non-parents (vs. with parents)	-0.23	0.21	-0.64	1.03	0.79	0.28		0.29	0.41	-0.51	1.09	1.34	0.47	
Years in Education	-0.08	0.03	-0.14	-0.19	0.92	0.09	*	-0.15	0.08	-0.31	0.01	0.86	0.048	*
Cannabis (vs. no cannabis)	-0.08	0.18	-0.43	0.77	0.92	0.67		-0.05	0.31	-0.66	0.56	0.95	0.88	
Non-smoker (vs. smoker)	-0.24	0.20	-0.63	1.00	0.79	0.24		0.03	0.35	-0.66	0.72	1.03	0.93	
Non-drinker (vs. drinker)	-0.09	0.21	-0.50	0.89	0.91	0.66		-0.06	0.45	-0.94	0.82	0.94	0.90	
PANSS Positive	0.03	0.02	-0.01	0.05	1.03	0.09		0.03	0.04	-0.05	0.11	1.03	0.46	
PANSS Negative	0.07	0.02	0.03	0.01	1.07	0.001	**	0.08	0.05	-0.02	0.18	1.08	0.12	
PANSS General Psychopathology	0.04	0.01	0.02	0.00	1.04	0.01	*	-0.01	0.04	-0.09	0.07	0.99	0.88	
SAPS	0.05	0.03	-0.01	0.07	1.05	0.08		0.05	0.07	-0.09	0.19	1.05	0.52	
SANS	0.02	0.01	0.00	0.02	1.02	0.14		-0.03	0.04	-0.11	0.05	0.97	0.51	
GAF	-0.02	0.01	-0.04	0.06	0.98	0.004	**	-0.03	0.02	-0.07	0.01	0.97	0.07	
No qualifications (vs. reference)	-0.42	0.22	-0.85	1.25	0.66	0.06		-0.29	0.46	-1.19	0.61	0.75	0.54	
Basic education qualification (vs. reference)	-0.05	0.21	-0.46	0.85	0.95	0.79		-0.33	0.40	-1.11	0.45	0.72	0.42	
Further education qualification (reference)	NA							NA						
Higher education qualification (vs. reference)	-1.11	0.25	-1.60	2.03	0.33	< 0.001	***	0.18	0.46	-0.72	1.08	1.20	0.70	
Non-European (reference)	NA							NA						
Black (vs. reference)	0.40	0.15	0.11	0.19	1.49	0.01	*	0.18	0.49	-0.78	1.14	1.20	0.71	

Asian/Other (vs. reference)	0.01	0.24	-0.46	0.91	1.01	0.98	-0.27	0.37	-1.00	0.46	0.76	0.48

[1] Adjusted for all predictors included in the model, as well as length of follow-up and cohorts.
NB: *, p < 0.05; **, p < 0.01; *** p < 0.001
Abbreviations: see List of Abbreviations on page 15; 95%CI, 95% confidence interval; SE, standard error; NA, not applicable.

Table 10. Coefficients for the LASSO logistic regression.

	Apparent Coefficients		**Recalibrated Coefficients**	
Predictor	**Coefficient**	**Direction of Effect**	**Coefficient**	**Direction of Effect**
Intercept	0.32		0.26	
Age of onset (years)	-0.02	-	-0.02	-
DUP (days)				
Female (vs. male)	-0.10	-	-0.09	-
Body Mass Index	0.002	+	0.002	+
Being in a relationship (vs. single)	-0.15	-	-0.14	-
Living alone (vs. with others)				
Living with non-parents (vs. with parents)	0.06	+	0.06	+
Years in Education	-0.04	-	-0.03	-
Cannabis (vs. no cannabis)				
Non-smoker (vs. smoker)				
Non-drinker (vs. drinker)	-0.33	-	-0.31	-
PANSS Positive				
PANSS Negative				
PANSS General Psychopathology				
SAPS				
SANS				
GAF	-0.02	-	-0.02	-
No qualifications (vs. all other qualifications)				
Basic education qualification (vs. all other qualifications)				
Further education qualification (vs. all other qualifications)	0.16	+	0.15	+
Higher education qualification (vs. all other qualifications)				
Non-European (vs. European)				
Black (vs. non-black)				
Asian/Other (vs. non-Asian/Other)				

Table 11. Performance measures for the LASSO logistic regression.

	Apparent Performance Measures	**Corrected Performance Measures**
Performance Measure	**19.1% Threshold**	**19.1% Threshold**
Alpha	0.47	-0.05
Beta	1.32	0.96
AUC	0.65	0.60
Accuracy	0.64	0.57
Sensitivity (%)	58.87	58.88
Specificity (%)	64.72	64.71
Positive Predictive Value (%)	26.01	24.13
Negative Predictive Value (%)	88.19	88.47

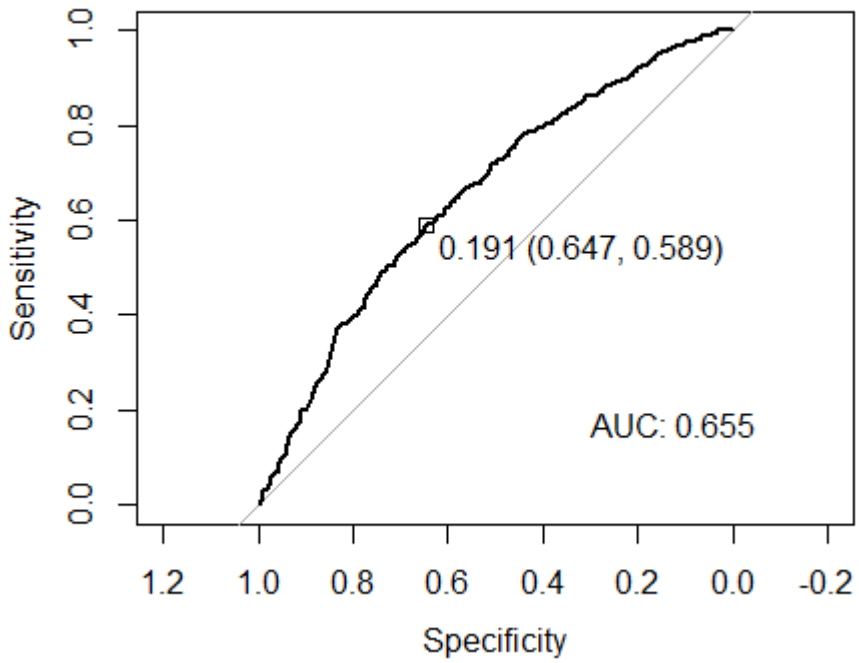


Figure 2. Area under the receiver operating characteristic curve (AUC) of the LASSO regression using the 1SE Model. To maximise sensitivity and specificity, the best threshold would be 0.191.

9.4. Discussion

The results of this study show that treatment resistance can be predicted with modest accuracy, using only clinical and demographic measures, when a patient first presents to clinical services, at the start of their illness. The model, corrected

for optimism, was able to identify 59% of TR patients and of 65% NTR. 24% of TR patients were correctly classified while 89% of NTR were correctly classified. This is the largest analysis of TR conducted to date; repurposing data from many research groups and driven by an international collaboration.

In my explanatory model, for a one-year decrease in age of onset there was an increase in the odds of being TR. Previous research has found that younger age of onset is associated with both TR (Wimberley et al., 2016b) and a broad range of poor outcomes in schizophrenia (Immonen, Jääskeläinen, Korpela, & Miettunen, 2017). Immonen et al. (2017) postulate that younger age of onset results in poor outcomes because it disrupts social and cognitive development. In a meta-analysis, Rajji, Ismail, and Mulsant (2009) found that adult and adolescent-onset schizophrenia were both associated with cognitive deficits in multiple domains, but that those with adolescent-onset schizophrenia had greater deficits on domains including IQ, executive functioning, and verbal memory. Adolescent-onset psychosis has also been associated with more negative symptoms at first-presentation (Downs et al., 2019) and worse premorbid functioning between the ages of 16 and 18 (Ballageer, Malla, Manchanda, Takhar, & Haricharan, 2005). Cognitive deficits, particularly poor verbal memory (de Bartolomeis et al., 2013; Joober et al., 2002), an increased burden of negative symptoms (Iasevoli et al., 2018b), and worse premorbid functioning (Albert et al., 2011; Levine & Rabinowitz, 2010) have all been associated with nonresponse to antipsychotic treatment. It is therefore unclear whether these impairments are the result of an earlier age of onset or characteristic of a treatment resistance subtype of schizophrenia. In my sample, younger age of onset, worse functioning, fewer years in education, and more negative symptoms were all significantly associated with TR, but only younger age of onset and fewer years in education remained so in the multivariable explanatory model. This suggests that these latter two may be independently associated with TR. However, it must be noted that I did not control for any cognitive variables. In contrast, all these features were selected in the predictive model, suggesting that, regardless of their true causal nature, they all contribute statistical power to predict new unseen cases.

In attempting to predict remission in psychosis patients two years after first episode, Emsley et al. (2006), using support vector machines, were able to identify 89% of remitters and 86% of non-remitters using only demographic and clinical information available at first episode. Although this prediction model was tested on hold-out data, the sample sizes of both the training and test data were very small (N=25 and N=16, respectively) and may not be generalisable. While many studies have examined predictors of first-line antipsychotic response at first episode (Suvisaari et al., 2018; Zhu et al., 2017) or predictors of clozapine response within patients with TR (Samanaite et al., 2018), I know of only two studies that have used prediction modelling to identify predictors of TR at first episode (Demjaha et al., 2017; Lally et al., 2016a). In a sample of first episode psychosis patients from south London, who were followed up after ten years, Demjaha et al. (2017) used a penalised regression model to predict TR. (Lally et al., 2016a) also use a penalised regression model to predict TR in a different south London sample of first episode patients who were followed up after five years. Only in Demjaha et al. (2017) was the statistical analysis optimised for prediction, with the use a LASSO regression. Lally et al. (2016a) used stepwise methods for variable selection based on the overall F value; a method which optimises explanatory power rather than predictive power (Shmueli, 2010). Neither study reported R^2 for their multivariable regressions. Demjaha et al. (2017) reported a Brier score of .146, which does not suggest good predictive accuracy given that the incidence of TR is 23% in their sample (a Brier score of 0 would be a perfect model while a score of .177 would be a non-informative model; Steyerberg et al., 2010). Lally et al. (2016a) reported no measure of predictive accuracy. While both studies had strong explanatory power, they lacked predictive power. Subsamples of data from these studies was included in the STRATA-G sample analysed in the current study.

TR has been thought by some as a biologically distinct subcategory of schizophrenia, and consequently there has been an active search for biomarkers (Gillespie et al., 2017). Despite – or perhaps because of – this, I have been able to predict TR with reasonable accuracy using only clinical and demographic features. Ascertainment of clinical information is less invasive, less time-consuming, and cheaper than biological markers. Without further validation, my findings cannot be used in clinical practice, but I suggest that future studies testing the predictive

accuracy of biological markers should also test whether the inclusion of clinical features identified in the present study can improve on the accuracy of a biomarker-only model. Alternatively, these clinical features may act as proxies for the underlying biology of the disease, and may be as equally effective in predicting TR as a biomarker-only model. Thus, a comparison of both types of models within the same sample is needed.

One of the strengths of this project is that I have estimated, separately, the strength of associations, between characteristics at first episode and TR, and the predictive accuracy of these same characteristics (Shmueli, 2010). My explanatory model uses data collected at first episode; many studies rely on case-control or retrospective designs and I cannot assume that findings from these studies will generalise to first episode patients. First episode data – and defining TR using prospective data – is especially important in prediction modelling. A LASSO regression is an optimal method for prediction in this context. Variable selection is an automatic part of the model, which produces a parsimonious model (Kuhn & Johnson, 2013); models with fewer predictors are easier to apply to new samples, since the amount of information required from participants is reduced, as well as easier to apply to existing samples, as there is a higher chance that all the required data has been collected. LASSO regression also prevents noise variables from having excess importance. This is especially important in my analysis since a limitation of the Random Forest missing data imputation method is that it can give more weight to noise variables with a large proportion of missingness, when the percentage of missingness varies across predictors (Lu & Petkova, 2014).

The main limitation of my statistical methodology is that I have not been able to externally validate my prediction model; this model, while having good internal validity, may not be able to predict new unseen cases in a different sample. Without external validation, I can make no claims about the clinical utility of this model. I would welcome any collaboration that could rectify this limitation and encourage other researchers to test this model on their own data. LASSO handles multicollinearity well, a consistent occurrence in clinical samples, and I took the further step of removing highly correlated variables before any analyses. LASSO regressions have been shown to work well in clinical data sets (Kuhn & Johnson,

2013), and have good predictive accuracy when determining psychosis onset (Ciarleglio et al., 2018), and schizophrenia diagnosis (Salvador et al., 2017). This model has yet to be externally validated, but it is not without merit: data was collected prospectively and only first episode variables, collected before illness duration or antipsychotic medication could influence this data, were included in the analysis.

Indeed, nearly all the variables selected by the model have previously been associated with TR, using explanatory models in longitudinal data: younger age of onset (Wimberley et al., 2016b), being male (Szymanski et al., 1995), being single (never married; Emsley et al., 2006), having more than a primary level of education (Wimberley et al., 2016b), and lower GAF score (worse functioning) (Horsdal, Wimberley, Kohler-Forsberg, Baandrup, & Gasse, 2018). Previous studies, when comparing TR patients and NTR, have failed to find a difference in the number of years in education (Frydecka et al., 2016; Iasevoli et al., 2018c), alcohol use (Chen, Chen, Chiu, Tai, & Lung, 2018), BMI (one study of treatment non-response found a difference in BMI at baseline; Chiliza, Asmal, Kilian, Phahladira, & Emsley, 2015), and who patients live with; either these are noise variable or previous studies have not been adequately power to detect an association, but regardless they add something to the predictive power of the model. However, the choice of variables was limited: as a result of using legacy data, variable collection was not informed by the current literature. More literature-informed variables may be needed to improve the AUC.

There are two key limitations of my participant sample: this is a sample of first episode *psychosis* patients – not *schizophrenia* patients – and the rates of TR vary between the cohorts. I did not restrict the main analysis to those with a schizophrenia diagnosis, because a recent meta-analysis suggested that patients with schizophrenia are often diagnosed with other psychotic disorders at first episode (Fusar-Poli et al., 2016). My analysis using only patients whose last known diagnosis at a follow-up visit was schizophrenia, had a substantially smaller sample size. Despite this, eight variables were selected by this schizophrenia-only model, five of which were also in the original model, and the AUC, sensitivity, specificity, PPV, and NPV were comparable to the original model. The differing rates of TR

reflect one of the statistical problems of international research consortia (Budin-Ljøsne et al., 2014). Due to the proportion of missing data in my sample, I was not able to use methods such as leave-one-out cross validation or meta-analyses to estimate the heterogeneity. However, Bühlmann and van de Geer (2018) argue that sample heterogeneity can produce more robust estimates in causal modelling and while heterogeneity is a problem for robust prediction, my model still had reasonable predictive accuracy in this worst-case scenario.

9.4.1. Conclusion

Our results show that younger age of onset and less years in education at first episode is associated with TR in a sample of psychosis patients. In addition, TR can be predicted with modest predictive accuracy using only clinical and demographic information recorded at first episode. At present, this is the largest sample of psychosis patients where TR status is known. Further work is required to test the predictive power of this model in other samples and to improve prediction accuracy by using datasets with additional variables (e.g. biological markers).

A GENOME-WIDE ASSOCIATION STUDY OF TREATMENT RESISTANCE

10. The Genetics of Treatment Resistance: A genome-wide association study

10.1. Summary

This chapter details work led by Dr Antonio Pardiñas which I contributed to. His draft manuscript is reproduced in Appendix D with his permission. The following chapter is my summary of this work.

10.1.1. Introduction

Given that only one second-generation antipsychotic, clozapine, has been shown to be efficacious for treatment resistant psychosis (TR) (Siskind, McCartney, Goldschlager, & Kisely, 2016; Taylor, 2017), it has been suggested the underlying biological mechanism of TR is different to that of non-treatment resistant psychosis (NTR). We hypothesise that the differences observed between TR and NTR patients may be the result of two distinct underlying genetic architectures. To date, the most influential studies examining the genetics of psychosis have failed to take into account treatment response. Nevertheless, antipsychotic treatment response-related loci have been identified. The Psychiatric Genetics Consortium (PGC) published the results of their genome-wide association study (GWAS) and identified 179 single nucleotide polymorphisms (SNPs) associated with schizophrenia, at 145 independent loci (Pardiñas et al., 2018; Ripke et al., 2014). Included in these SNPs were genetic risk loci that are located within the DRD2 gene. This gene encodes the dopamine receptor, which is targeted by antipsychotics; occupancy of the dopamine receptor is associated with the alleviation of psychotic symptoms. Risk loci, located within other neurotransmitter systems, were also identified. These may provide novel targets for treatment and/or identify biological causal mechanisms of schizophrenia (e.g. the GRIN2A gene in the glutamatergic system and the CACNA1C gene involved in neuronal calcium signalling). Therefore, investigating the genetics of TR, specifically, provides one possible route to establishing which biological systems underpin TR and identifying new treatment targets.

As well as revealing causal mechanisms, genetic data can also be used to predict health-related outcomes. The use of genetic predictors in clinical practice is seen positively by patients, their families, and health professionals (Jones, Scourfield, McCandless, & Craddock, 2002; Lawrence & Appelbaum, 2011). In research,

polygenic risk score (PRS) profiling is a statistical tool that is used to predict genetic risk for complex diseases. On a practical level, a PRS for treatment resistance could be an ideal predictor. It is cheap and relatively non-invasive to collect DNA; which is also stable from birth and consequently, unlike clinical predictors, not affected by confounding environmental factors other than ancestry (Lewis & Vassos, 2017). Although at present, PRS profiling lacks predictive power and is not used clinically in psychiatry, it is conceivable that for disorders with high heritability and heterogeneity, such as psychosis, some form of PRS may be developed for predicting prognosis in a clinical context (Palk, Dalvie, de Vries, Martin, & Stein, 2019).

10.1.2. Genome-wide association study (GWAS)

As briefly discussed in my introductory chapter, GWAS are a statistical method used to identify the common genetic risk factors of a disease. By comparing the frequency of SNPs in patients with TR to patients with NTR, we can identify differences in the genetic architecture between the two groups. GWAS have replaced family studies of genetics in examining diseases with a low prevalence in the population, such as schizophrenia and TR. However, they require large samples to detect the small effect that an individual SNP has on a disease. Large sample sizes are also necessary to counterbalance the stricter significance thresholds applied to compensate for multiple testing.

We conducted two GWAS analyses and one interaction analysis. The first GWAS used data from the Psychiatric Genetics Consortium to compare patients with schizophrenia to healthy controls (HC) (Ripke et al., 2014). Patients with known TR were excluded from the patient group so only NTR patients remained. In this GWAS there was a total sample size of N=48,798 (21,264 NTR and 27,534 HC). The second GWAS used data from CLOZUK1 and CLOZUK2, which consist of all patients prescribed clozapine in the UK, and compared them to HC from publicly available resources (Pardiñas et al., 2018). This GWAS had a total sample size of N=34,043 (10,501 TR and 24,542 HC). To compare TR and NTR patients we conducted a ‘virtual’ GWAS using a meta-analytic technique called a ‘test for interaction’ (Altman & Bland, 2003). For every SNP that was present in both the PGC and CLOZUK GWAS, we calculated the difference between the two odds

ratios, in the form of a z-score and its associated P-value. This method is more powerful than simply comparing the TR and NTR patients, as statistical power is increased by including the large HC samples.

In our analysis, one SNP was significantly associated with TR at the genome-wide significance level. This SNP, rs79780963, is an intronic SNP in the protein-coding NT5C2 gene, but has also been described as an expression quantitative trait locus (eQTL) – a genomic locus that explains the variance in expression of mRNA – in the ARL3 gene. rs79780963 has not been identified in any previous GWAS, but the NT5C2 gene has been linked with schizophrenia, autism, smoking, body mass index, and educational attainment (Buniello et al., 2018). NT5C2 was associated with schizophrenia in an early analysis of the PGC data (Ripke et al., 2011), the CLOZUK data (Pardiñas et al., 2018), as well as in Japanese, Ashkenazi Jew, Chinese, and Swedish samples (Bergen et al., 2012; Goes et al., 2015; Ikeda et al., 2019; Li et al., 2017).

However, the association between a single SNP and TR must be interpreted cautiously. Within the genome, SNPs are often in linkage disequilibrium (LD). SNPs are said to be in LD if they are inherited together more often than would be expected by chance; SNPs in LD often lie nearby to one another on the same chromosome (Pettersson et al., 2009). As a result of LD, a single SNP – with a causal effect on TR – would lead to multiple statistical, but non-causal, associations with nearby SNPs (Berisa & Pickrell, 2016). There are no SNPs nearby to rs79780963 which are also associated with TR, either above or below the genome-wide significance level. This can be seen in Figure 11 of Pardiñas et al (section 0). This may suggest that the association between rs79780963 and TR is a false positive finding. Alternatively, the SNPs in LD with rs79780963 may have similar effects sizes in the PGC and CLOZUK GWAS. The test of interaction we used is limited in that it is only able to detect large differences in effect size.

To draw stronger conclusions from these results, the findings need to be replicated in independent samples, and we chose to do this using polygenic risk score (PRS) profiling.

10.1.3. Polygenic risk score analysis

A PRS is the weighted sum of SNPs associated with a disease, where the association reaches a certain statistical threshold. SNPs were weighted by effect size in an additive model (Clarke et al., 2011). For this analysis we calculated three PRS, one using the effect sizes from the interaction analysis (TR-PRS), a second using the effect sizes from the clozapine vs healthy controls GWAS (CLOZ-PRS) and the final PRS using the effect sizes from NTR vs healthy controls GWAS (PGC-PRS). Each PRS was adjusted for principal components (PCs) and tested over eight P-value thresholds. PCs represent population stratification – variation in the genome due to differing ancestries – which may be associated with both SNPs and the disease in question, and, therefore, covarying for PCs reduces confounding-by-ancestry (Zondervan & Cardon, 2007). We tested all three PRS in two independent samples and compared TR to NTR patients.

CardiffCOGS is a cross-sectional psychosis cohort based on contemporaneous records (Legge et al., 2019; Pardiñas et al., 2018; Rees et al., 2014). In our analysis we used N=790 individuals from CardiffCOGS (TR = 341 (43%) and NTR = 449 (57%)). STRATA-G is comprised of first episode psychosis cohorts with at least one year follow-up; in our analysis we used N=1059 individuals from STRATA-G (TR = 148 (14%) and NTR = 911 (86%)) who provided DNA and whose data survived quality control (quality control was described in Chapter 7).

In both CardiffCOGS and STRATA-G, a higher TR-PRS was associated with a higher probability of being treatment resistant. In the CardiffCOGS sample, the TR-PRS explained 2.70% of the variance in TR status; the CLOZ-PRS explained 0.97%, whilst the PGC-PRS was not associated with TR. In the STRATA-G sample, the TR-PRS explained 1.11% of the variance, the CLOZ-PRS explained 1.49%, and again, the PGC-PRS was not associated with TR. This suggests that our test of interaction did indeed identify a genetic architecture specific to treatment resistance.

10.1.4. Genetic co-localisation analysis

We also tested the genetic correlations of TR and other traits, using LD score regression. An LD score is the measure of how many nearby variants a particular

SNP is co-inherited with. Specifically, it is the sum of squared correlations between the focal SNP and all other SNPs (Zheng et al., 2017). The LD score is used as a measure of how likely a SNP is to tag neighbouring SNPs that also affect the phenotype in question. In LD score regression, for each genetic variant, the LD score is regressed onto the GWAS test statistic. SNPs with a higher test statistic should also have a higher LD score, as they have a greater likelihood of being co-inherited with true causal variants.

This method is used to distinguish SNPs with high test statistics due to genuine polygenicity from SNPs with inflated test statistics due to population stratification. LD score regression is used as a measure of heritability when examining one trait – as the genetic correlation between a trait and itself is equivalent to narrow-sense heritability – but it can also be used to examine genetic correlation between two different traits. In this scenario, the LD score is regressed onto the product of the two test statistics derived from the two GWAS examining the different traits. LD score regression does not require individual-level data and instead uses GWAS summary statistics; plus, it is not biased by sample overlap between the two GWAS. Genetic correlation uses all SNPs, regardless of whether they pass the genome-wide significance threshold (for more information on LD score regression and genetic correlation see Bulik-Sullivan et al., 2015a; Bulik-Sullivan et al., 2015b; Zhang et al., 2018; Zheng et al., 2017).

We applied false discovery rate (FDR) correction for multiple testing. Ever having smoked and the number of cigarettes smoked per day were positively correlated with TR. Educational attainment, having a university degree, adult and childhood IQ, being a former smoker, and having a diagnosis of bipolar disorder were all negatively correlated with TR. Only educational attainment, having a university degree, adult IQ, and being a former smoker were statistically significant. These correlations must be interpreted with caution as they may be spurious associations, rather than true signals, given the low heritability of the TR polygenic signal.

When two traits are genetically correlated, the SNPs which contribute to those traits are usually inherited together. To examine this idea further, we used the significant correlations in co-localisation analysis. Two traits are said to be co-localised if the

two independent associations (at the same locus), generated by the two separate GWAS, share a causal variant. Evidence of co-localisation increases the probability that two traits share a causal mechanism (Giambartolomei et al., 2014). As expected, we found there was a negative genetic correlation between cognitive phenotypes and TR; poorer educational attainment and adult IQ correlated with a higher risk of TR, with 42 co-localised SNPs. We also found that smoking behaviour was positively genetically correlated with TR, with 2 co-localised SNPs.

10.1.5. Conclusion

Our work suggests that common genetics variants contribute to the genetic architecture of TR. The polygenic risk score for treatment resistance (TR-PRS) captures a small proportion of the genetic heritability of TR, and while individuals with higher TR-PRS are more likely to be TR, it remains to be seen whether this can be a useful predictor in clinical settings. A GWAS cannot capture all the complexity in the genetic architecture of a disease, such as pleiotropy – multiple phenotypic effects of single variants – so we sought to address this by conducting genetic co-localisation analysis. The genetic overlap between TR and cognition that we identified replicates findings from non-biological studies of TR. But the genetic overlap with smoking behaviour is an interesting finding given that anecdotally TR patients are more likely to smoke, but previous studies have not tested this statistically (Iasevoli et al., 2013; Kennedy et al., 2014).

AGE OF ONSET: THE PATH FROM POLYGENIC RISK TO TREATMENT RESISTANCE

11. Age of Psychosis Onset: The path from polygenic risk to treatment resistance

11.1. Introduction

Within psychosis spectrum disorders, the age at which an individual's psychotic symptoms first occur has been identified as a prognostic factor of long-term health outcomes. Younger age of psychosis onset (age of onset) has been associated with poor outcomes including relapse in the first year after onset and a second hospitalisation (Immonen *et al.*, 2017). Younger age of onset has also been repeatedly associated with antipsychotic treatment resistance (Demjaha et al., 2017; Lally et al., 2016a; Meltzer et al., 1997; Wimberley et al., 2016b). Treatment resistance (TR) is defined as the persistence of psychotic symptoms after two trials of antipsychotic medication, trials that have lasted for six weeks once a therapeutic dose has been prescribed. TR is a serious concern in the treatment of people with psychosis, since patients with TR have a poorer quality of life than those who respond to antipsychotics (Iasevoli et al., 2016) and their care and treatment is particularly costly for healthcare services (Andrews et al., 2012).

Age of psychosis onset is itself heritable. Using a maximum likelihood method and data from 717 patients in 327 families, Hare et al. (2010) estimated the heritability of age of onset to be 33%. This was after the authors had controlled for sex, which accounted for 2% of the variance in age of onset. In addition, younger age of onset has been linked to increased genetic susceptibility for developing schizophrenia. A national database study from Denmark found that a second twin was 4.69 times more likely to develop schizophrenia if the first twin developed schizophrenia *before* the age of 22 (Hilker et al., 2017). In Sweden, a national database study found that the risk associated with developing schizophrenia, if a parent had psychosis, was stronger in patients with a younger age of onset (Li, Sundquist, & Sundquist, 2007). A meta-analysis of 15 studies found that family history of schizophrenia was associated with younger age of onset, an association that could not be explained by sex, the degree of relation, or severity of symptoms (Esterberg, Trotman, Holtzman, Compton, & Walker, 2010). This suggests that an increased familial genetic load for schizophrenia results in psychosis occurring earlier in an individual's lifetime. Psychosis that develops in adolescence or early twenties may

indicate a more severe psychosis phenotype and a more heritable form of the disorder.

To date, I am aware of only one study that tested whether the relationship between age of onset and TR could be explained by genetic liability to schizophrenia. Legge et al. (2019) reported that the association between age of onset and TR was not attenuated by the polygenic risk score for schizophrenia. While this study controlled to some extent for genetic susceptibility to schizophrenia, it did not control for genetic susceptibility to TR. Furthermore, none of the previous studies that have linked younger age of onset to TR controlled for genetic susceptibility for TR.

My aim was to test the hypothesis that the association between age of onset and treatment resistance is confounded by an increased genetic load for TR. I assumed that there is a shared genetic load between schizophrenia and treatment resistance, because studies of the genetic load for schizophrenia have not excluded individuals with TR. There is a strong argument that those with a greater genetic load for schizophrenia have an earlier psychosis onset, therefore it is likely that individuals with a greater genetics load for treatment resistance will also have an earlier psychosis onset. In order to test this, I investigated the pathways between genetic susceptibility for TR, age of psychosis onset, and treatment resistance. To investigate whether the relationship between age of onset and treatment resistance is confounded by an increased genetic load *specifically* for TR, I separately tested whether genetic load for non-treatment resistant schizophrenia, educational attainment, and IQ also acted as confounders. Given the genetic correlation between treatment resistance and both educational attainment and IQ, it is highly likely that the genetic load associated with these traits may also confound the relationship between age of onset and treatment resistance. There has also been considerable success in using polygenic risk scores for cross-trait prediction (Krapohl et al., 2016), including one study which reported that an increased polygenic risk score for schizophrenia was associated with earlier onset of depression (Power et al., 2017).

11.2. Methods

11.2.1. Participants

STRATA-G is a sample of treatment resistant (TR) and non-treatment resistant individuals (NTR) with psychosis, created by combining observational, prospective, first episode psychosis cohorts with a minimum of one year follow-up. For this analysis, I used a subsample of STRATA-G participants who had provided DNA, and whose genotypes passed quality control thresholds, and for whom age of onset was known.

11.2.2. Treatment resistance

Criteria for TR has been already determined for other analyses. I used three definitions of TR, depending on the data available in each cohort: either (1) prescription of or treatment with clozapine, or (2) treatment with two different antipsychotics at a therapeutic dose for a specified duration (e.g. each of at least 6 weeks' duration, with dosages in at least the mid-point of the licensed therapeutic range), or (3) persistent psychotic symptoms (e.g. as rated by the PANSS) and moderate functional impairment (e.g. as rated by the GAF) despite treatment with two different antipsychotics at a therapeutic dose for a specified duration of time. Any participant not meeting criteria was considered to be NTR (for more detail on how TR was defined see Chapter 7).

11.2.3. Age of onset

Age of onset was considered to be the participant's age when psychotic symptoms first occurred. If this variable was not available, the participant's age when they first presented to clinical services for psychosis was used. As date of first presentation to clinical services is likely to be systematically later than date of first psychotic symptoms, I applied a correction to age at first presentation to clinical services. One cohort included in this sample (the AESOP study) collected data on both age at first presentation and age of first symptoms. The mean difference between these ages was 0.547 years. Therefore, the estimated age of onset, in the absence of the variable, equalled age at first presentation to clinical services minus 0.547 years.

11.2.4. Genotypes

Called genotypes underwent quality control, as described in the Chapter 7, prior to imputation (homozygosity rate of the X chromosome: ≥ 0.8 for males and ≤ 0.2 for females; single nucleotide polymorphism (SNP) call rate > 95%; missingness < 5%; identity-by-state (IBS) > 0.9). Imputation was performed on subsamples of STRATA-G separately, rather than on the whole sample together, as this is thought to be a more robust method to maximise the overlap in SNPs than imputing all the samples together (Stanaway et al., 2019). SNPs were imputed using the Michigan Imputation Server and the Haplotype Reference Consortium panel (McCarthy et al., 2016). Post-imputation filters were applied to the genotypes (INFO > 0.8; Probability threshold > 0.9; Missingness < 5%; Hardy Weinberg Equilibrium P-value > 10^{-6}).

11.2.5. Polygenic risk scores

Five polygenic risk scores (PRS) were calculated and taken forward into subsequent analyses. Three PRS for treatment resistance, clozapine use, and non-treatment resistant schizophrenia (TR-PRS, CLOZ-PRS and PGC-PRS, respectively) were calculated using the summary statistics from the two GWAS and one interaction test described in Chapter 10. The CLOZ-PRS was computed using a genome-wide association study (GWAS), which compared individuals who were treated with clozapine to healthy controls. The PGC-PRS was computed using a GWAS that compared patients with schizophrenia, excluding those known to be treatment resistant, to healthy controls. The TR-PRS was calculated using the difference in SNP effect sizes between these two GWAS. The polygenic risk score for education attainment (EA-PRS) was calculated using the summary statistics reported in Lee et al. (2018) and for IQ (IQ-PRS) using Savage et al. (2018). The SNPs associated with educational attainment and adult intelligence were shown to be co-inherited with the SNPs associated with treatment resistance in work by Dr Pardiñas (Chapter 10). All five GWAS controlled for principle components (PCs) (following procedures described in Chapter 10, Lee et al. (2018), Pardiñas et al. (2018), Ripke et al. (2014), and Savage et al. (2018)).

As described in Chapter 7, PRS were calculated using LDpred-inf (Vilhjalmsson et al., 2015). It has been suggested that the power of polygenic risk scores can

increase with the number of included SNPs (Dudbridge, 2013). The predictive accuracy can also be increased when markers in linkage disequilibrium are included (Vilhjalmsson et al., 2015).

To account for the effect of ancestry on the polygenic risk scores, PRS in the first imputation batch were adjusted for the first 5 PCs, as well as principle component 10 and 11, while in the second batch PRS were adjusted for only the first 5 PCs. Linear regressions were used to calculate the residual polygenic risk scores, adjusted for the PCs.

11.2.6. Statistical Analysis

Figure 3. shows the path models that were tested in the following analyses.

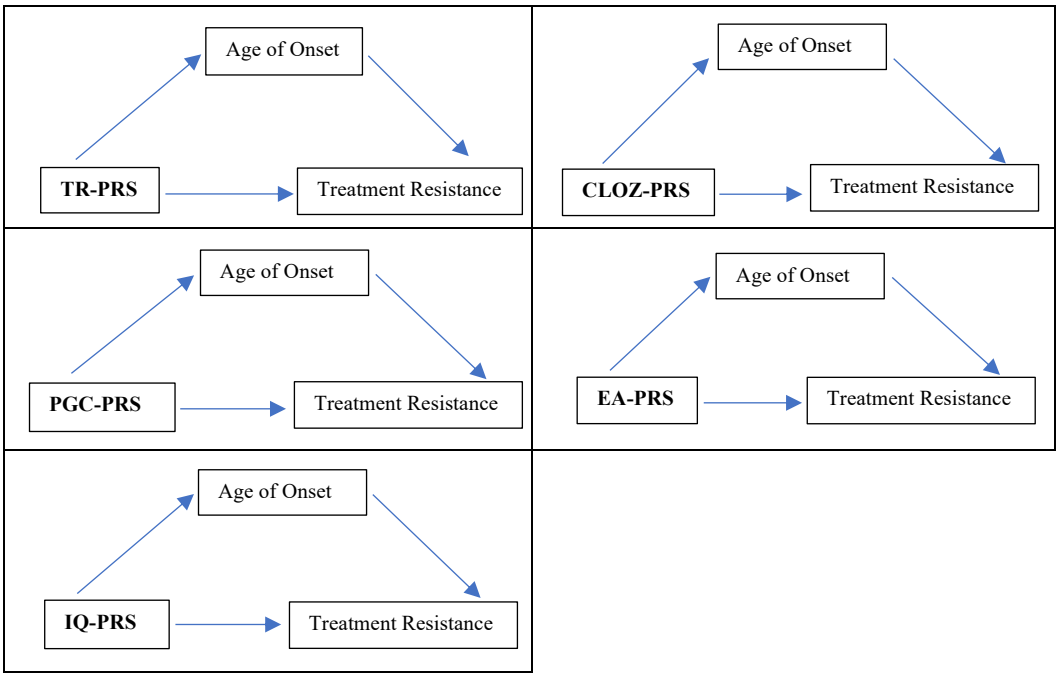


Figure 3. Input diagram showing the five models and the unidirectional associations between variables which were tested. Abbreviations: TR-PRS, polygenic risk score for treatment resistance; CLOZ-PRS, polygenic risk score for clozapine use; PGC-PRS, polygenic risk score for non-treatment resistant schizophrenia; EA-PRS, polygenic risk score for education attainment; IQ-PRS, polygenic risk score for IQ.

11.2.7. Age of onset and treatment resistance

As age of onset is highly skewed (skew = 1.05, kurtosis = 0.84), Mann Whitney *U* tests were performed to compare age of psychosis onset in the TR and non-treatment resistant (NTR) groups. The mean and standard deviation of age of onset was stratified by cohort and imputation batch, then meta-analysed using a random effects model and the restricted maximum-likelihood method.

11.2.8. Age of onset and polygenic risk scores

Partial Spearman's correlations were calculated to compare age of onset and the five polygenic risk scores (calculated using LDpred-inf). Rather than use the residual of the PRS, PCs were controlled for by computing partial correlation coefficients. After stratification by cohort and imputation batch, these partial correlation coefficients were meta-analysed, using a random effects model and the restricted maximum-likelihood method.

11.2.9. Polygenic risk scores and treatment resistance

Independent t-tests were used to compare the difference in residual polygenic risk scores in the TR and NTR groups. The mean and standard deviation of age of onset was stratified by cohort and imputation batch, then meta-analysed using a random effects model and the restricted maximum-likelihood method.

11.2.10. Polygenic risk scores, age of onset, and treatment resistance

Path analysis was conducted in R using the 'lavaan' package. Unlike other packages, lavaan can incorporate endogenous variables that are categorical. Only individuals with non-missing data on gender were entered into the path analysis i.e. N = 789 (85%) of the total sample (N = 925).

Assuming the hypothesised causal relationships, I estimated the direct effect between the polygenic risk score for treatment resistance and TR, the indirect effect via age of onset, and the total effect. I repeated this analysis four more times: replacing the polygenic risk score for treatment resistance with the polygenic risk scores for clozapine, non-treatment resistant schizophrenia, educational attainment, and IQ (

Figure 3). Causal relationships were assumed to be unidirectional and time-ordered. A robust diagonally weighted least-squares estimation, rather than an asymptotically distribution-free estimator, was used as this method works well with smaller sample sizes but also non-normally distributed data (Rhemtulla, Brosseau-Liard, & Savalei, 2012). Bootstrapped standard errors are reported. The models were assessed using the chi-squared statistic, which measures the difference between the expected and observed covariance matrices.

When using the chi-squared test to estimate model fit, the risk of type I errors increases when variables have non-normal distributions. Therefore, I also report three fit indices as recommended by Kenny (2015) and use these to evaluate model fit. The Root Mean Square Error of Approximation (RMSEA) is an absolute measure of fit, which uses the ratio between the chi-squared statistic to the degrees of freedom to penalise the model for complexity. An absolute measure of fit presumes that the best fitting model has a fit of zero. Hu and Bentler (1999) suggests that an acceptable model fit is indicated by RMSEA value less than 0.06. MacCallum, Browne, and Sugawara (1996) have used 0.01, 0.05, and 0.08 to indicate excellent, good, and mediocre fit, respectively. The Comparative Fit Index (CFI) is an incremental measure of fit based on the non-centrality measure and pays a penalty for every parameter estimated. A value of 0 indicates the worst possible model fit and a value of 1 indicates the best possible model. If the index is greater than one, it is set at one and if less than zero, it is set to zero. The CFI depends on the average size of pairwise correlations between variables in the data. If the average correlation between variables is not high, then the CFI will be low. The Akaike Information Criterion (AIC) is a comparative measure of fit and is therefore only interpretable when comparing multiple different models. Lower values indicate a better fit and so the model with the lowest AIC is the best fitting model. AIC was calculated using the following formula $\chi^2 + k(k + 1) - 2df$ where k is the number of variables in the model and df is the degrees of freedom.

11.2.11. Gender

Evidence suggests that males have a younger age of onset (Eranti, MacCabe, Bundy, & Murray, 2013), therefore a two-way Analysis of Variance (ANOVA) was used to examine the interaction between gender and treatment resistant status

on age of psychosis onset. Path analysis models were repeated stratifying on gender.

11.2.12. Power Analysis

In structural equation modelling, a sample size of N=200 is usually required with a ratio of 20 participants for each free parameter (Kenny, 2015). Each model contained three variables and five free parameters (I did not estimate the disturbance associated with TR), therefore, a minimum of 100 TR cases are required for the analysis to be adequately powered when the outcome is categorical variable (Streiner, 2005).

11.2.13. Thresholding

The alpha level was set at 0.05. As genotypes were imputed in two imputation batches, results were meta-analysed when possible. When meta-analysis was not possible, the second imputation batch was treated as a replication sample. I give more weight to findings which were replicated.

All analyses were performed in R version 3.6.0 (R Core Team, 2009) and the MAJOR package in *jamovi* (Version 1.0.2.0) was used to perform meta-analyses.

11.3. Results

11.3.1. Sample Size

There were N=1061 individuals in STRATA-G with genotype data that passed all genotype quality control thresholds. N=135 individuals were then excluded because their age of psychosis onset was unknown and N=1 individual was excluded as they were not of European ancestry. The final analytic sample contained N=925 (TR=135 (14.59%)) participants (Table 12).

After splitting by imputation batch, in the first batch there was N=708 individuals of which 95 (13.42%) were TR. In the second imputation batch, there were N=217 individuals of which 40 (18.43%) were TR. The number of participants stratified by cohort is reported in Appendix E (Table 29).

Table 12. Descriptive statistics of participants with genotypes that survived quality control, stratified by treatment response and cohort.

	Total	NTR	TR	Statistic	P-value
N (%)	925	790 (85.41)	135 (14.59)	NA	
Mean age of onset (SD)	26.78 (8.79)	27.21 (8.93)	24.23 (7.50)	NA	
Median age of onset (IQR)	24.69 (20.00-31.58)	25.1 (20.50-32.50)	21.5 (19.00-28.40)	855625[1]	< 0.001
N of males (%)	457 (57.34)	374 (44.76)	83 (69.17)	7.5[2]	0.01
Mean TR-PRS (SE)	$-4.101e^{-17}$ (0.07)	-0.01 (0.08)	0.08 (0.17)	-0.49[3]	0.60
Mean CLOZ-PRS (SE)	$4.647e^{-17}$ (0.65)	-0.50 (0.70)	2.91 (1.67)	-1.90[3]	0.06
Mean PGC-PRS (SE)	$-7.09e^{-17}$ (0.70)	-0.48 (0.76)	2.78 (1.87)	-1.60[3]	0.10
Mean EA-PRS (SE)	$3.973e^{-18}$ (0.01)	0.004 (0.01)	-0.02 (0.02)	1.1[3]	0.30
Mean IQ-PRS (SE)	$-5.029e^{-16}$ (1.51)	1.17 (1.63)	-6.74 (3.93)	1.9[3]	0.06

NB. The polygenic risk scores here correspond to the means and standard errors (SE) of the residuals of the polygenic risk scores; [1] Mann–Whitney U statistic; [2] Chi-squared statistic [3] Independent samples t-test statistic.

11.3.2. Age of onset and treatment resistance

Participants with TR had a younger age of psychosis onset. The median age of onset was 21.5 in the TR patients and 25.1 in the NTR patients; the distributions in the two groups differed significantly (Mann–Whitney U = 855625, P < 0.001). Means and standard deviations are reported in Table 12 and, when stratified by cohort and imputation batch, in the Appendix E (Table 30).

When meta-analysed, TR participants had a younger mean age of onset than NTR participants by 2.77 years (standard error = 0.61, 95% CI: 3.97, 1.58) (Figure 4). In terms of heterogeneity, Tau^2 = 0.65 (standard error = 1.51, P = 0.35) and I^2 = 18.17% (Figure 5).

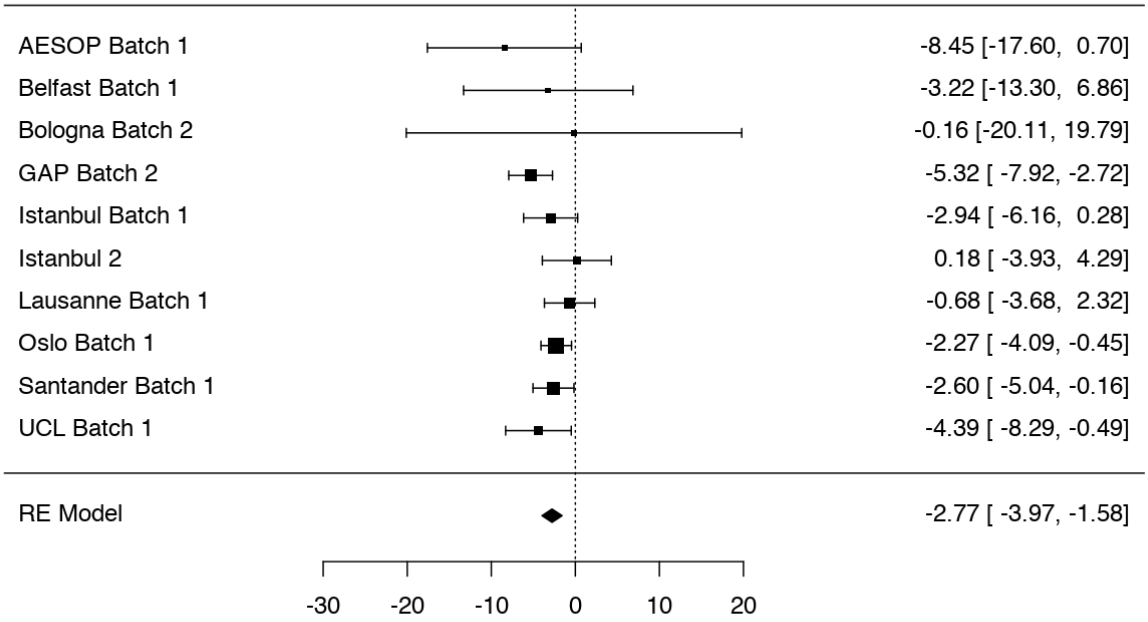


Figure 4. Forest plot of the raw mean difference in age of psychosis onset between treatment resistant and non-treatment resistant participants.

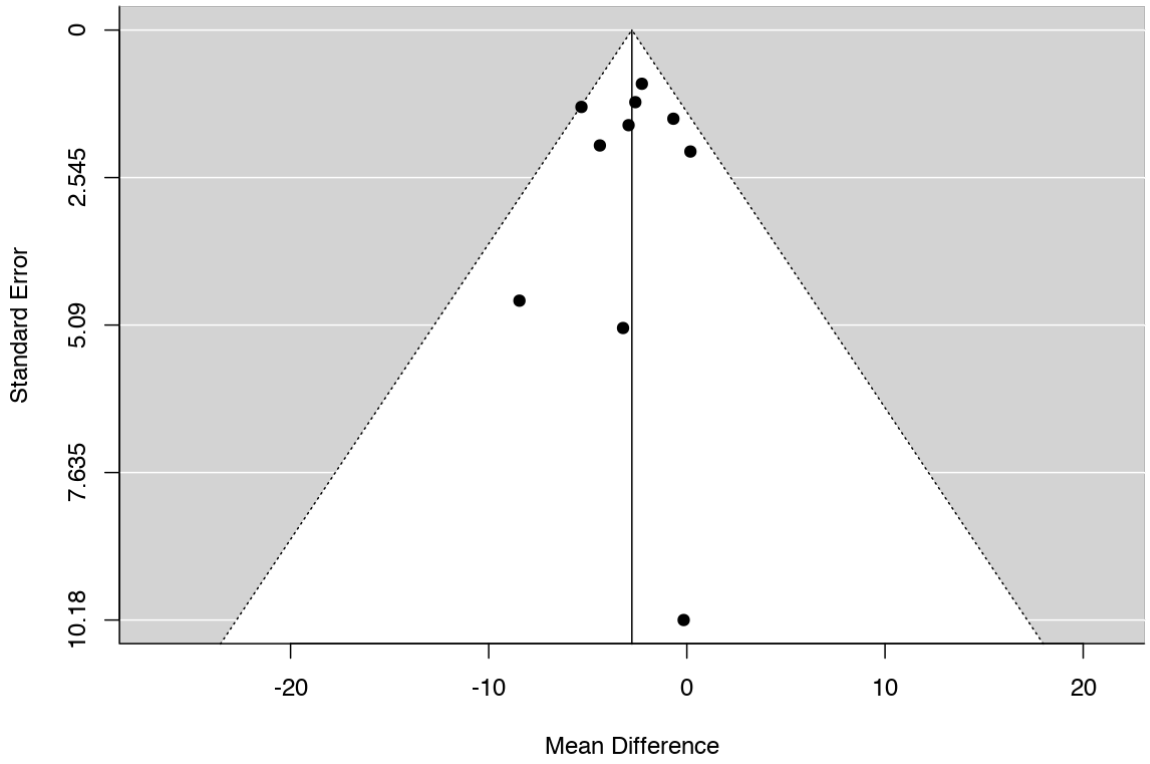


Figure 5. Funnel plot of the raw mean difference in age of psychosis onset between treatment resistant and non-treatment resistant participants.

11.3.3. Age of onset and polygenic risk scores

Age of psychosis onset was not significantly correlated with the TR-PRS, CLOZ-PRS, PGC-PRS, EA-PRS, or IQ-PRS in either of the imputation batches (Table 13). Partial correlations, stratified by treatment resistance status, cohort, and imputation batch, are reported in Appendix E (Table 31).

When meta-analysed, there was no significant correlation between age of onset and TR-PRS (r = -0.01, standard error = 0.04, 95%CI: -0.10, 0.07) or the clozapine polygenic risk score (r = -0.02, standard error = 0.03, 95% CI = -0.08 – 0.05). There was also no significant correlation between age of onset and non-treatment resistant polygenic risk score (r = 0.00, standard error = 0.03, 95% CI = -0.06 – 0.07). In terms of cognition, there was no significant correlation between age of onset and the educational attainment polygenic risk score (r = 0.00, standard error = 0.03, 95% CI = -0.06 – 0.07) or the IQ-PRS (r = 0.01, standard error = 0.03, 95% CI = -0.06 – 0.07). Additional information on heterogeneity is reported in the Appendix E (Table 32), alongside forest and funnel plots for all five meta-analyses (Appendix E Figure 17 to Figure 26).

When stratifying by treatment resistance status (Table 13), the PGC-PRS was positively correlated with age of onset in treatment resistant participants in the second imputation batch (r = 0.35, P = 0.04). Treatment resistant participants with an older age of onset had a higher PGC-PRS (the PRS based on summary statistics comparing NTR schizophrenia patients vs. healthy controls). However, this was not observed in the first imputation batch and was only nominally significant (i.e. it would not exceed any correction for multiple testing).

Table 13. Partial correlation coefficients for age of psychosis onset and the five polygenic risk scores calculated using LDpred-inf.

	Imputation Batch 1						**Imputation Batch 2**					
	TR (N=95)		**NTR** (N=603)		**All** (N=698)		**TR** (N=40)		**NTR** (N=177)		**All** (N=217)	
	r	P-value	r	P-value	r	P-value	r	P-value	r	P-value	r	P-value
TR-PRS	-0.09	0.43	0.02	0.59	0.01	0.89	-0.30	0.08	-0.00	1.00	-0.08	0.23
CLOZ-PRS	-0.10	0.35	-0.01	0.86	-0.03	0.39	-0.06	0.75	-0.05	0.49	-0.06	0.42
PGC-PRS	-0.04	0.72	-0.01	0.82	-0.02	0.53	0.35	0.04*	-0.07	0.35	0.04	0.55
EA-PRS	-0.10	0.34	0.03	0.49	0.02	0.60	-0.03	0.88	-0.06	0.42	-0.02	0.77
IQ-PRS	-0.08	0.48	-0.03	0.48	-0.03	0.51	-0.21	0.22	0.01	0.89	-0.01	0.85

NB. Estimates in Imputation Batch 1 are adjusted for the first 5 principle components and principle components 10 and 11. Estimates in Imputation Batch 2 are adjusted for the first 5 principle components.
*, P < 0.05; **, P < 0.01; *** P < 0.001

11.3.4. Polygenic risk scores and treatment resistance

The polygenic risk scores based on treatment resistance, clozapine, and non-treatment resistance schizophrenia did not differ significantly between TR and NTR patients. Meta-analyses did not show a significant effect. There was also no difference in PRS for educational attainment and IQ between TR and NTR patients, nor did meta-analyses show a significant difference between the two groups. Means and standard deviations are reported in Table 12. See Appendix E for means and standard deviations stratified by cohort (Table 33) and the results of the meta-analyses (Table 34 as well as Figure 27 to Figure 36).

11.3.5. Path analysis: polygenic risk scores, age of onset, and treatment resistance

N=789 participants, of which TR=120 (15.21%) and NTR= 669 (84.79%), were entered into the path models. Although all five models were statistically different from the null, all the models were a poor fit of the data according to both the RMSEA and CFI (Table 14). Using the AIC, the model including the clozapine polygenic risk score was the model with the best comparative fit.

Table 14. Model fit statistics for path analysis.

	TR-PRS	CLOZ-PRS	PGC-PRS	EA-PRS	IQ-PRS
Chi-squared statistic	4337.62	4333.21	4334.69	4334.77	4333.47
Chi-squared P-value	< .001	< .001	< .001	< .001	< .001
Degrees of freedom	3	3	3	3	3
Root Mean Square Error of Approximation (RMSEA; 90%CI)	1.35 (1.32, 1.39)	1.35 (1.32, 1.39)	1.35 (1.32, 1.39)	1.35 (1.32, 1.39)	1.35 (1.32, 1.39)
Comparative Fit Index (CFI)	0	0	0	0	0
Akaike Information Criterion (AIC)	4343.62	4339.21	4340.69	4340.77	4339.47
R^2 age of onset	0%	0%	0.1%	0%	0%
R^2 treatment resistance status	5.2%	5.8%	5.7%	5.7%	6%

In terms of specific paths, as expected, age of onset had a direct effect on TR in all five models. The direct, indirect, and total effects of the polygenic risk score for TR on treatment resistance are shown in Figure 6. Output diagrams for the other four models are shown in the Appendix E (Figure 37, Figure 38, Figure 39, and Figure 40).

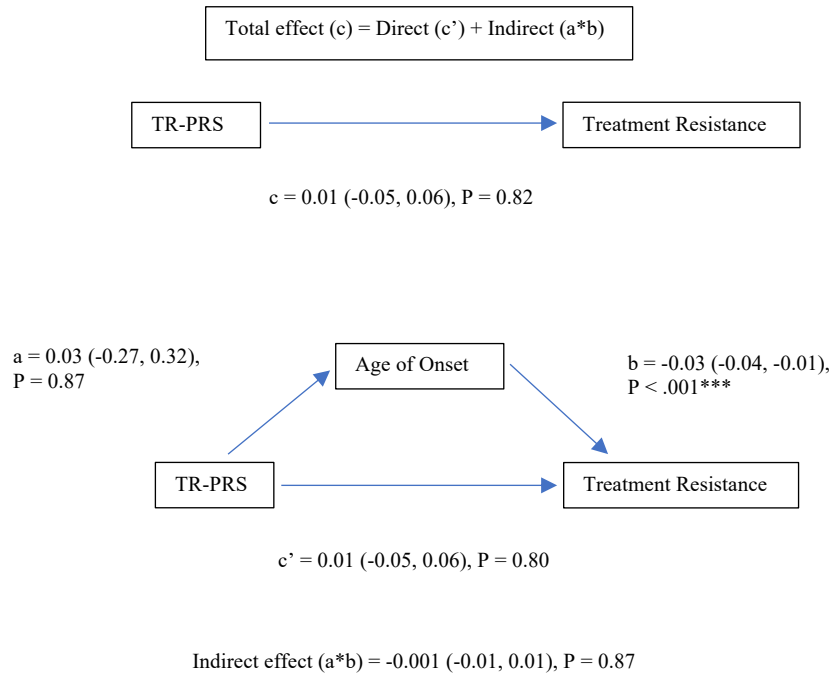


Figure 6. Output diagram showing the direct effect between the polygenic risk score for treatment resistance and treatment resistance, and the indirect effect with age of psychosis onset as a mediator.

11.3.6. Gender

Male participants have a younger age of psychosis onset (male mean = 25.25 ± 7.40 vs. female mean = 29.92 ± 10.51) and TR participants have a younger age of onset (TR mean = 24.23 ± 7.50 vs. NTR mean = 27.21 ± 8.93). The distribution of age of onset is shown in Figure 7 and Figure 8.

There was a statistically significant interaction between the effects of gender and treatment resistance on age of onset ($F(1, 793) = 4.46$, $P = 0.035$; see Figure 9). Tukey's HSD post hoc tests showed that within females, TR participants had younger age of onset than NTR participants ($P = 0.002$), but within males there was no difference in age of onset between TR and NTR participants ($P = 0.50$). Within NTR participants, males had a younger age of onset than females $P < 0.001$), but within TR participants there was no difference in age of onset between male and female participants ($P = 0.93$). There was no difference in age of onset between male NTR participants and female TR participants ($P = 0.99$), but male TR participants had a younger age of onset than NTR female participants ($P < 0.001$).

When the five path analysis models were stratified by gender, the direct effect of younger age of onset on TR was only significant in women.

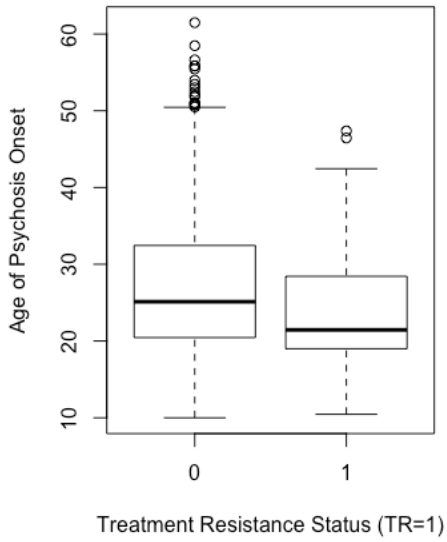


Figure 7. Box-and-whisker plot of age of psychosis onset stratified by treatment resistance status.

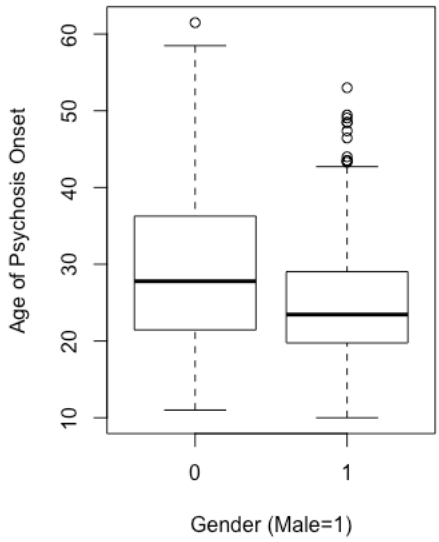


Figure 8. Box-and-whisker plot of age of psychosis onset stratified by gender.

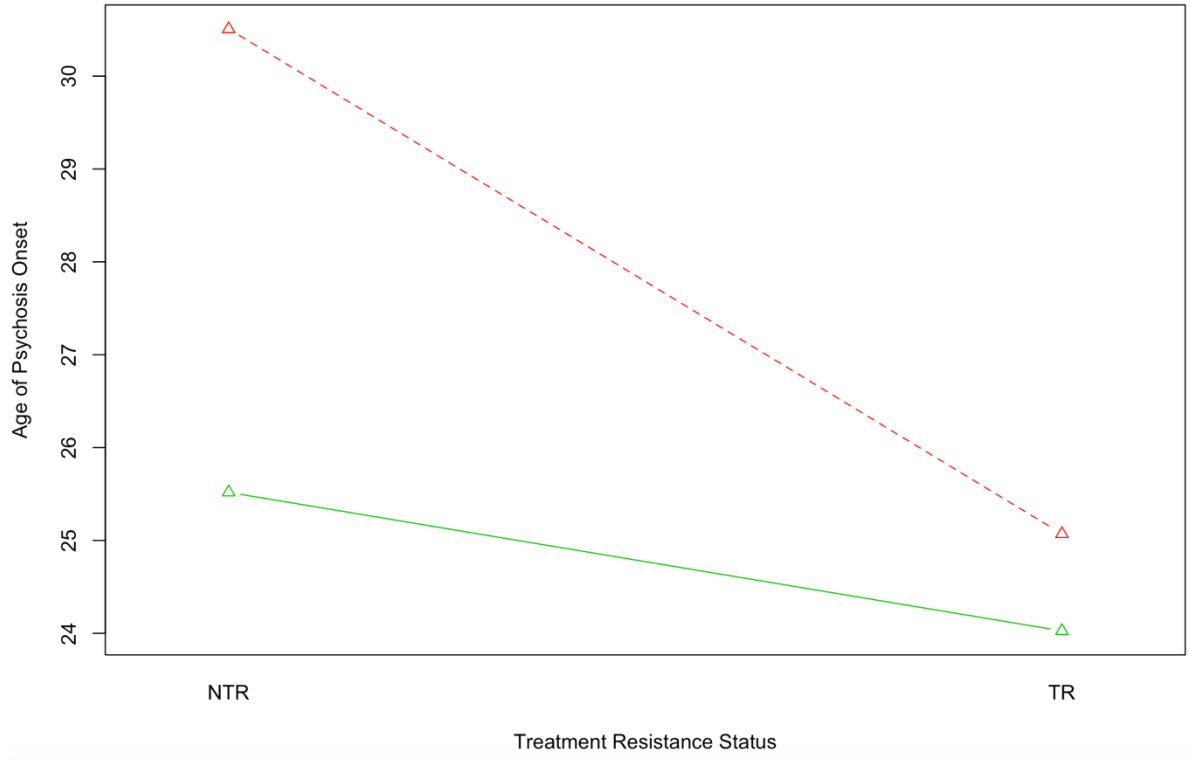


Figure 9. Plot of the interaction between treatment resistance status and gender on age of psychosis onset (male = green line; female = red line; TR = treatment resistant; NTR = non-treatment resistant; age of onset is measured in years).

11.4. Discussion

In STRATA-G, I found that TR participants had a younger age of psychosis onset than NTR participants by a mean of 2.77 years. Neither polygenic risk scores for treatment resistance nor the polygenic risk scores for clozapine use, non-treatment resistant schizophrenia, educational attainment, or adult IQ were associated with treatment resistance. I did not find any association between age of psychosis onset and these polygenic risk scores, that was replicated across the two imputation batches or was identified in a path model with good fit. This suggests that the association between having a younger age of psychosis onset and later developing treatment resistance may not be due to a higher genetic liability for either treatment resistance, as hypothesised. Nor is it due to a higher genetic liability for clozapine use, non-treatment resistance or other co-localised genetic traits. These data therefore support a previous study which reported that the genetic liability for schizophrenia did not account for the association between younger age of psychosis onset and treatment resistance (Legge et al., 2019).

Men are consistently shown to have a younger age of psychosis onset than women, with one study estimating that males have a younger age of onset than females by approximately 1.49 years (Eranti et al., 2013). Although polygenic risk scores are calculated using the autosomes, and therefore cannot be confounded by sex, gender may still confound the relationship between genetic susceptibility for a trait and age of onset. Esterberg et al. (2010) found that males had a younger age of onset only when patients had no family history of psychosis; there was no difference in age of onset between males and females when adjusting for family history of psychosis. Hilker et al. (2017) report that the association between younger age of onset and the likelihood that a second twin will also develop schizophrenia is larger in females than males. This suggests that higher genetic risk for schizophrenia is more influential on lowering age of psychosis onset in women than in men. Although at least one study has found no difference in the relationship between genetic risk and age of onset, after stratifying by sex (Li et al., 2007), in my data, I have found evidence that younger age of onset appears to have a greater impact on subsequent treatment resistance in women than in men. Furthermore, I found that, once stratified by gender, the path between younger age of onset and treatment resistance was observed in women only.

Tentative findings from this study suggest that a higher polygenic risk score for non-treatment resistant participants is correlated with an older age of psychosis onset, but only in treatment resistant participants. Because the effects of the polygenic risk scores on TR is not consistent across subgroups, the polygenic risk score for non-treatment resistance could be used as predictor of TR if it was combined with other variables. Prediction requires heterogeneity and variables which capture patterns in the data can be combined to form a prediction model (Steyerberg, 2008). This is even true for variables that are not causally associated with an outcome, as they can still improve the accuracy of prediction models (Steyerberg, 2008).

11.4.1. Strengths & Limitations

Limitations in the design of STRATA-G have been discussed in Chapter 8 and considered further in Chapter 12. There is variation in the prevalence of TR across

the individual cohorts which make up the total sample, but I have attempted to account for this by analysing the studies separately and then using meta-analysis where possible. Despite the biases inherent in observational longitudinal designs, STRATA-G is the largest first episode cohort where treatment resistance status is known. In the path analysis, there were 24 TR cases per parameter, which is more than the suggested 20 cases per parameter required for an adequately powered analysis (Streiner, 2005).

In terms of the path analysis, a major limitation was that the principle components, which capture genetic variation attributable to geographical location and ethnic background, could not be included in the models directly. Despite the large sample size there was insufficient statistical power to include seven more variables. However, employing a two-step approach, obtaining PRS residuals adjusted for the principal components, should be a good approximation to a model which included PCs as covariates. Nevertheless, it is possible that this alternative approach does not fully capture the confounding effect that the principal components may have on the relationship between genetics of treatment resistance status and age of onset.

There is also considerable debate about the correct use of fit indices. For example, Barrett (2007) argues that only the chi-squared statistic should be interpreted, while Hayduk, Cummings, Boadu, Pazderka-Robinson, and Boulianne (2007) argue that cut-offs for a fit index are misleading and should not be used at all. Part of the problem lies in using fit indices to specify that a model is ‘good’ as opposed to ‘bad’. Path analysis cannot be used to determine whether a specific model is correct, it can only be used to determine whether the data are consistent with the model or not (Streiner, 2005). Kenny (2015) argues in favour of reporting fit indices but not relying on strict cut-offs. In terms of my models, all are a good fit of the data according to the chi-squared statistic. However, this may be a false positive finding attributable to either the large sample size (relative to what is typically used in the structural equation modelling literature) or the skewed distribution of age of onset. Therefore, in using the RMSEA and the CFI, I can conclude that none of the proposed models are a good fit of my data. Age of onset may mediate the effect of SNPs – captured by the polygenic risk score for treatment

resistance – on TR, but the STRATA-G data do not provide significant evidence in favour of this hypothesis.

Another problem in path analysis is that models can be 'underidentified', that is, that the number of parameters to be estimated in the model is greater than the number of parameters that it is possible to estimate (Streiner, 2005). To avoid this, I did not estimate the disturbance associated with treatment resistance i.e. the error which captures the imprecision in how treatment resistance is defined and the effect of confounding variables on treatment resistance.

I have conducted a number of statistical tests to fully explore the relationship between genetic load for treatment resistance, age of onset, and subsequent treatment resistance. I did not apply a correction for multiple testing as the statistical tests reflected a combination of primary and secondary hypotheses and were not independent of one another (e.g. considering the correlation between the five PRS). This makes it highly challenging to estimate an appropriate level of significance. Introducing a correction for multiple testing could increase the risk of type II errors. Instead, unadjusted P-values were provided throughout and should be interpreted cautiously, accounting for the strength of prior evidence for any associations tested and the range of tests performed. For instance, the probability of finding an effect size with the degree of magnitude observed between age of onset and treatment resistance reported here, if there were no such effect in the underlying population, is unlikely to be due to sampling variation.

11.4.2. Conclusion

In conclusion, my data have replicated the finding that younger age of psychosis onset is associated with treatment resistance. Despite evidence suggesting that age of onset is itself heritable, these data do not reliably suggest that age of onset is associated with genetic risk for treatment resistance, educational attainment, or IQ or that genetic load can account for why age of onset is consistently associated with treatment resistance.

GENERAL DISUCSSION

12. General Discussion

12.1. Summary of findings

There is a clinical need to identify individuals with antipsychotic treatment resistant psychosis, earlier in the course of their illness. Treatment resistant (TR) patients have a worse prognosis than their non-treatment resistant counterparts (NTR); they suffer more from impaired social and occupational functioning and experience a poorer quality of life (Brain et al., 2018; Iasevoli et al., 2016; Nordstroem et al., 2017). Furthermore, there is evidence across psychosis spectrum disorders that a shorter period between symptom onset and treatment (Murru & Carpiniello, 2018) – and a shorter delay between symptom onset and successful treatment (Malla et al., 2006) – is associated with a better prognosis.

At present, TR patients can only be identified after they have tried and failed to respond to two separate antipsychotics, when both are prescribed at a therapeutic dose and taken for at least six weeks. Identifying those likely to be TR when they initially present to clinical services for psychosis, would have the greatest impact on reducing delays to successful treatment. Using variables collected at initial presentation also removes the potentially confounding effect of antipsychotic medication. The aim of this thesis was to identify predictors of treatment resistant psychosis, using data collected at this point; from patients when they presented to clinical services for their first episode of psychosis.

12.1.1. Predictors of Treatment Resistant Schizophrenia: A systematic review of prospective observational studies

First, I reported the results of a systematic review of research examining predictors of TR (Chapter 8). I focused exclusively on prospective observational studies where baseline data was collected at either first episode or first treatment with antipsychotics. In clinical practice, misclassification of patients as TR is common due to poor treatment compliance and subtherapeutic antipsychotic dose or plasma levels. This bias is therefore also present in observational studies which rely on data collected from routine clinical records. However, this means that, compared to patients who agree to participate in clinical trials, patients in observational studies are more representative of psychosis patients who enter clinical services.

Therefore, I considered predictors identified from observational studies to be more robust to confounders in clinical practice and have greater generalisability.

I identified eleven published studies, and one unpublished study, which used data from eight distinct longitudinal cohorts. Age of psychosis onset was the most tested risk factor for TR. Seven out of the eight studies measured age of onset and six found a significant association between younger age of onset and TR (Chan et al., 2014; Demjaha et al., 2017; Kim et al., 2017; Lally et al., 2016a; Meltzer et al., 1997; Wimberley et al., 2016b). Four studies tested gender; two reported that males were more likely to be TR than NTR (Lally et al., 2016a; Meltzer et al., 1997). Three studies examined duration of untreated psychosis and one of these found TR patients had a longer duration of untreated psychosis than NTR (Demjaha et al., 2017). Of the three studies that examined season of birth, two reported an association between being born in Autumn/Winter and TR (Kim et al., 2017; Sorensen et al., 2014). All other risk factors were only tested in one or two of the cohorts.

Two main gaps in the literature were identified. First, there was a marked lack of studies investigating biological predictors of treatment resistance; only two studies included a biological risk factor. Wimberley et al. (2017) tested the polygenic risk score for schizophrenia and Horsdal, Wimberley, Benros, and Gasse (2017a) tested C-reactive protein levels but neither were associated with TR. Second, none of the studies used analytic methods optimised for prediction. Most studies used either logistic regressions or Cox logistic regressions to estimate the magnitude of effect between risk factors and TR status. Only Lally et al. (2016a) and Demjaha et al. (2017) used penalised logistic regressions, a method designed to reduce overfitting, which is better than non-penalised regressions at predicting outcomes in new data. Furthermore, only Demjaha et al. (2017) and Wimberley et al. (2016b) reported performance measures for prediction (McFadden's pseudo R^2 and Harrell's C statistic, respectively). None of the studies internally validated their models, using either bootstrapping or cross-validation, or tested their models in holdout or external data. Nor did any of these studies perform unbiased variable selection by incorporating this process as an automatic, internal part of the model development.

I attempted to address the lack of biological predictors and robust prediction modelling in the rest of this thesis.

12.1.2. Clinical and Demographic Predictors Treatment Resistance: A prediction model

I next reported the results of my own data analysis, conducted with the aim of identifying predictors of TR (Chapter 9). Using clinical and demographic data collected at first episode I created a model that had an area under the curve of 0.60. The area under the curve is the probability that the model assigns a higher probability of being TR to a TR patient than a NTR patient (Steyerberg, 2008). An uninformative model will have an area under the curve of 0.5, while a perfect model will have an area under the curve of 1. It is a useful method of predictive accuracy as it is independent of the occurrence of TR.

I used a LASSO logistic regression that selected nine variables: TR was predicted by younger age of onset, male gender, higher BMI, being single, living with someone other than their parents, fewer years in education, drinking alcohol, lower GAF score (poorer functioning), and achieving a higher education qualification. This model was able to correctly classify 59% of TR patients and 65% of NTR patients. However, there was also considerable misclassification in my model. Of all the patients classified as TR, only 24% were actually TR, and of all the patients classified as NTR, only 88% were actually NTR. This may be, in part, due to the low prevalence of TR in the sample (17%). As a consequence, in a hypothetical situation where the model is used to choose a clinical intervention, any treatment for TR should not be determinantal to the NTR patients misclassified as TR. I discuss other uses of this prediction model in section 12.3.1.

Nevertheless, a model that accurately classified 59% of TR patients and 65% of NTR patients would be more cost effective than current treatment practice in the UK. Jin et al. (2019) calculated the cost of implementing a model, such as mine, in clinical practice. They considered a test that could be used after patients had failed to respond to one antipsychotic. Those who were unlikely to respond to a second antipsychotic, according to the test, would be offered clozapine as a second-line treatment. Compared to treatment as usual (i.e. clozapine as a third-line treatment

for all patients), and taking into consideration the side effects of first-line antipsychotics and clozapine, using a predictive model in this scenario would result in improvement of 0.10 quality-adjusted life years (QALY) and reduce healthcare costs by £7363 per person. Their analysis suggests that a predictive model would be more cost effective than treatment as usual, even if it only identified 6% of TR patients and 50% of NTR patients. Although, for many clinicians and patients, a model with such poor classification would not be considered a viable option for clinical practice.

In examining the results of my prediction model, it is important to note that even though a variable is selected in the model, this variable may not be associated with, or be a causal factor of, treatment resistance. Prediction models will include spurious associations alongside 'true' associations, if inclusion of those variables increases the predictive accuracy of the overall model. Neither should the direction or magnitude of effect between each predictor and TR be used to make claims about the nature of TR. These associations have all been penalised to avoid overfitting the sample data. It was for these reasons that I also conducted a logistic regression, which did not penalise the coefficients, to provide a so-called explanatory model of TR. When I examined each variable separately, younger age of onset, male gender, fewer years in education, higher PANSS negative score, lower GAF score (worse functioning), achieving a higher education degree, and being black were all significantly associated with TR. In the prediction model, four of these variables were selected, as they contributed to the predictive accuracy of the overall model. In the explanatory model, after controlling for all other variables, only younger age of onset and fewer years in education were associated with TR. In STRATA-G, for every one-year increase in age of psychosis onset, the odds of being TR decrease 0.95 times. Whilst for every one-year decrease in years of education, the odds of being TR decrease 0.86 times. I discuss age of onset and years in education later in sections 12.2.1 and 12.2.2, respectively.

12.1.3. The Genetics of Treatment Resistance: A genome-wide association study

Next, I reported the results of a genome-wide association analysis (GWAS) for treatment resistance (Chapter 10). Three separate polygenic risk scores were tested

against TR status in the STRATA-G sample. The PGC-PRS (non-treatment resistant psychosis patients vs. healthy controls) was not associated with TR, but the CLOZ-PRS (treatment resistant patients vs. healthy controls) explained 1.49% of the variance. The polygenic risk score for treatment resistance (TR-PRS) explained 1.11% of the variance. The proportion of variance explained is much smaller than proportion of variance in schizophrenia explained by the general schizophrenia polygenic risk score (Ripke et al., 2014), but comparable to other psychiatric polygenic risk scores (Krapohl et al., 2016). This suggests that there is a difference in the genetic architecture of TR and NTR psychosis. It may account for why, of the four previous studies which tested the general schizophrenia polygenic risk score against TR (a combined sample of TR and NTR patients vs. healthy controls), only two found that a higher polygenic risk score was significantly associated with TR (Frank et al., 2015a; Legge et al., 2019). In addition, we found that patients with a higher genetic load for treatment resistance also had a higher genetic load for poorer cognition, specifically fewer years in education and lower adult IQ (I discuss this further in section 12.2.2), as well as a higher genetic load for smoking. Previous genetic studies have introduced confounding by recruiting cases and controls (e.g. TR and NTR patients) from separate populations and genotyping cases and controls separately (Lambert & Black, 2012). As TR and NTR patients were recruited from the same population and also genotyping together, using STRATA-G as validation sample does not introduce these biases.

12.1.4. Age of Psychosis Onset: The path from polygenic risk to treatment resistance

I found younger age of onset to be repeatedly associated with TR. Furthermore, there is evidence to suggest that age of onset is itself heritable with one study estimating that the proportion of variance explained by genetics is 33%, after adjusted for gender (Hare et al., 2010). Therefore, I explored whether the genetics of treatment resistance could account for the observed relationship between age at psychosis onset and treatment resistance (Chapter 11). I hypothesised that the genetics of treatment resistance was a causal factor for both younger age of onset and being TR, and therefore age of onset was not a ‘true’ causal factor of TR. I used path analysis (a form of structural equation modelling sometimes referred to

as causal modelling) to measure the direct and indirect pathways between the variables. Younger age of onset was, as expected, associated with TR, but it did not act as mediator in the relationship between genetic liability for TR and subsequent treatment resistance. Here, genetic liability was measured using the polygenic risk score for treatment resistance. I also tested age of onset as a mediator of the relationship between the genetic liability for other traits and treatment resistance using polygenic risk scores for clozapine use, non-treatment resistance, educational attainment, and IQ. None of the polygenic risk scores were individually associated with TR status. However, after stratifying into subgroups, the polygenic risk score for non-treatment resistant participants was positively correlated with older age of psychosis onset specifically in treatment resistant participants. Age of onset may be influenced by genetics specific to antipsychotic treatment response, but only within certain subgroups of patients. An alternative explanation is that by stratifying into groups with smaller sample sizes, statistical power was reduced, and this finding was the product of sampling variability. Overall, no path model displayed a good fit to the data, and caution must be exercised in drawing conclusions about whether age of onset is associated with the biological mechanism that underpins treatment response.

12.2. Findings in context and implications

12.2.1. Younger age of onset and treatment resistance

The most replicated finding in my systematic review was that individuals with TR had a younger age of psychosis onset (Chapter 8). I replicated this finding myself using the STRATA-G data. I found that younger age of onset was automatically included in the prediction model designed to predict future cases of TR (Chapter 9). The association between age of onset and TR could not be explained by the polygenic risk score for treatment resistance (Chapter 11).

Younger age of onset may have a pleiotropic effect; not only is it associated with TR, but also with more hospitalisations, more negative symptoms, more relapses, as well as poorer social and occupational functioning (Immonen et al., 2017). However, all of these prognostic outcomes are also associated with TR, so it is unclear whether younger age of onset leads to a broad, more severe psychosis

phenotype or whether it is specific to TR (i.e. it is treatment resistance status that confounds the associations between younger age of onset and poor outcomes).

Age of psychosis onset itself may have several causes. One study found that patients with a history of cannabis use had a younger age of onset, but so did those with both a family history of psychosis and a history of obstetric complications (O'Donoghue et al., 2015). Tobacco use was also thought to be associated with age of onset, but this relationship has been shown to be confounded by cannabis use (Hickling et al., 2017). It is presumed that genetic and environmental factors together cause psychosis (Dean & Murray, 2005), for example, under a threshold model (Gottesman & Shields, 1972) or a diathesis-stress model (Pruessner, Cullen, Aas, & Walker, 2017). Under this assumption, the more risk factors for psychosis an individual has, the younger their age of psychosis onset is likely to be. In support of this theory, O'Donoghue et al. (2015) found that a higher number of environmental risk factors for psychosis were associated with a younger age of onset.

The association between age of onset and TR may suggest that, rather than a biologically distinct subgroup within psychosis, TR represents the severe end of the psychosis spectrum. This has implications for how to examine treatment resistance in research, as by defining TR categorically and using NTR as a comparison group, valuable information may be lost. One solution would be to categorise patients, after two trials of antipsychotic medication, into three groups: patients in remission, patients with treatment resistance, and patients who fall between these two categorical definitions. A similar approach is being used in schizophrenia research where patients with schizophrenia are compared to healthy controls and individuals with schizotypy – a cluster of personality traits, including bizarre behaviour, magical thinking, unusual perceptions, and social anhedonia which is similar to psychosis (Barrantes-Vidal, Grant, & Kwapil, 2015; Nelson, Seal, Pantelis, & Phillips, 2013). It is thought that for measurable characteristics of schizophrenia, when tested at a group level, individuals with schizotypy will fall between healthy controls and patients with schizophrenia. For treatment resistance, using three groups could facilitate a dimensional approach to understanding the aetiology of TR.

The finding that younger age of onset is associated with treatment resistance also has implications for clinicians, as patients who develop psychosis earlier in life (e.g. before the age of twenty; Lally et al., 2016a) will require closer monitoring for TR so that they do not experience a delay in receiving clozapine treatment.

12.2.2. Less years in education and treatment resistance

In my analysis, I found that fewer years in education was associated with TR and that, like age of psychosis onset, years in education was automatically included in the prediction model (Chapter 9). We also found that there is significant overlap in the genetics of treatment resistance and educational attainment (Chapter 10), even though the polygenic risk scores for educational attainment and IQ are not associated with treatment resistance (Chapter 11).

Previous studies have found that patients with TR have poorer cognitive abilities than NTR patients (de Bartolomeis et al., 2013; Frydecka et al., 2016; Joober et al., 2002; Kravariti et al., 2018; Legge et al., 2019), but none have tested an association with years in education, which could be considered a proxy measure for cognitive ability. Chan et al. (2014) did test years in education as a risk factor for TR but found no association (Chapter 8). While I found that fewer years in education was associated with TR in a regression models, this may not be the best statistical method to use since the association between the two variables may not be linear. When looking at highest educational qualification obtained before the first episode of psychosis (a categorical variable with four levels), it was completing A-levels (or the equivalent) that was associated with TR, and not a lack of qualifications, completing the equivalent to GCSEs, or obtaining a university degree. This contradicts findings from a Danish cohort study which showed that having a basic education qualification was associated with TR (Wimberley et al., 2016b), but here they split education qualification into only two categories.

When we compared the genetics associated with TR to the genetics associated with years of schooling we found a significant negative correlation (Chapter 10). SNPs associated with fewer years of schooling, were also associated with TR status. The years of schooling GWAS used a discovery cohort of 293,723 individuals, and a

replication cohort of 111,349 individuals (Okbay et al., 2016). As years of schooling may vary across countries due to differences in national policy, we also looked at college completion using a GWAS with 95,427 individuals in the discovery cohort and 25,490 in the replication sample (Rietveld et al., 2013). Again, SNPs associated with not completing college were also associated with being TR. It would require more research, but this supports a hypothesis that the cognitive impairments observed in cases of TR may not be a product of their psychosis, or the failure to treat their psychosis, but are in fact the result of polygenic effects that contribute to both poor cognition and TR. This could then explain why antipsychotics are less successful in treating cognitive impairments associated with schizophrenia (Takeuchi et al., 2017).

12.3. Future directions and implications

12.3.1. Prediction modelling in treatment resistance

I have produced the most comprehensive prediction model for treatment resistance to date. By incorporating automatic variable selection, penalising regression coefficients, applying an optimism correction, and evaluating the model using measures of predictive accuracy, I have overcome many of the limitations identified in the previous literature. Outside of national population registries, it is the largest first episode psychosis sample, which has been used to predict treatment resistance.

To fully test the predictive accuracy of my model, it will need to be tested in independent samples (the gold-standard method of validation). Only by examining the model's ability to discriminate between TR and NTR in new samples would I be able to definitively ascertain whether the model is better than chance at predicting TR. Measures of sensitivity and specificity of the model, in this new dataset, should then be used to ascertain whether the model would be beneficial in clinical practice and quantify the risk of patients being misclassified.

As well as testing the current model in new data it would also be necessary to consider additional predictors and assess how variable selection and predictive accuracy change accordingly. For example, we could re-run the existing model on the same subsample of STRATA-G data but also include polygenic risk scores for

treatment resistance or clozapine use. This would be viable as both variables explained a small proportion of variance in treatment resistance status and neither are correlated with age of psychosis onset. Even though the polygenic risk scores were not associated with treatment resistance status in explanatory models (Chapter 9), like BMI and alcohol use, they could still improve overall predictive accuracy of the model. Another option is to include multiple polygenic risk scores; a recent study showed that including multiple polygenic risk scores in a penalised regression explained more of the variance in outcome than the best polygenic risk score alone (Krapohl et al., 2018). This was true when predicting educational achievement, general cognitive ability, and BMI. Specifically, polygenic risk scores for traits in genetic correlation with treatment resistance (as identified in Chapter 10) could be used for predicting treatment resistance (Krapohl et al., 2018). Regardless of whether genetic correlation is due to pleiotropy or reflects a shared biological pathway, cross-trait prediction can help improve the accuracy of the prediction model, without requiring an understanding of the underlying mechanisms of TR.

There are two other variables – not available in STRATA-G – which have the potential to predict treatment resistance: brain glutamate levels and antipsychotic response in the first two weeks of treatment.

Demjaha et al. (2014) reported that TR was associated with higher levels of glutamate in the anterior cingulate cortex when compared to healthy controls; a difference that is not found when comparing NTR patients to healthy controls. Higher levels of glutamate in the anterior cingulate cortex have been found in TR patients and patients who failed to respond to only one antipsychotic trial (Egerton et al., 2012; Mouchlianitis et al., 2015). However, this finding is not consistent across all studies: Goldstein, Anderson, Pillai, Kydd, and Russell (2015) reported no differences in glutamate levels in the anterior cingulate cortex but instead found that patients with TR had higher glutamate/glutamine levels in the putamen. In the largest sample of chronic psychosis patients to date (N=92), Egerton et al. (Unpublished) found that anterior cingulate cortex glutamate levels, after adjusting for age and sex, were higher in TR patients compared NTR patients. This measure alone had an area under the curve of 0.59. Combing my clinical and demographic

predictors with a measure of glutamate may improve the model's ability to predict TR. However, the cost of 1H-magnetic resonance spectroscopy (the MRI technique used to quantify concentrations of metabolites in predefined neural regions) would have to be factored in when considering the clinical cost-effectiveness of a model that requires a measure of glutamate.

A second variable to consider is antipsychotic response in the first two weeks of treatment. A review by Carbon and Correll (2014) concluded that the presence, or absence, of minimal symptomatic improvement in the first four weeks of antipsychotic treatment is associated with response/nonresponse after a single-antipsychotic trial. This was the case in both samples of chronic and first episode patients. A meta-analysis of 34 single-antipsychotic trials found that a lack of response in the first two weeks of treatment predicted nonresponse at the study endpoint (between 4 to 12 weeks after medication was started) (Samara et al., 2015). Early response/nonresponse correctly classified 86% of nonresponders and 63% of responders (the corresponding positive and negative predictive values were 90% and 53%, respectively). In this meta-analysis early nonresponse was defined as <20% reduction of total PANSS or BPRS score from baseline to week two, while nonresponse at the study endpoint was defined as a <50% PANSS or BPRS reduction. Both Gillespie et al. (2017) and Bozzatello, Bellino, and Rocca (2019) conclude in their reviews that early nonresponse could be a predictor of TR, but neither cites work that has tested early nonresponse as predictor of TR, defined using categorical criteria or two antipsychotic trials. I know of no studies that have tested nonresponse in the first two weeks of treatment as a predictor of TR. However, a strategy using early nonresponse to determine whether patients should be switched to a second (non-clozapine) antipsychotic earlier in the illness course (a switch vs stay algorithm) has been successfully tested (Kinon et al., 2010). It is therefore possible that early nonresponse could be used to identify a subgroup of patients who would benefit from clozapine as second-line antipsychotic. As such a predictive model combining early nonresponse and my clinical and demographic predictors should be tested.

It is important to note that the use of these measures as predictive markers of TR is currently untested. Identifying potential predictors from studies that were designed

to estimate the magnitude of effect can lead to overoptimistic expectations of the markers predictive performance (Steyerberg et al., 2010). Nevertheless, additional predictors can be beneficial, even if they only have incremental value on top of predictors that are easily collected. How beneficial depends on the increase in predictive performance and the cost – in terms of money, time, and invasiveness to a patient – of collecting that particular predictor. Validation of my model, which contains readily accessible predictors, in fully independent data is the best way to compare the performance the model with, and without, a new predictor such as a polygenic risk score, glutamate level, or early nonresponse. In adding new predictors, it is important to remember that one of the key benefits of this model is that it allows risk to be determined at the *start* of antipsychotic treatment. The aim of the model is to prevent or mitigate the active psychotic symptoms that patients with TR experience before being successfully treated.

12.3.2. Clinical utility of predicting treatment resistance

I have identified potential and actual predictors of TR, but it was outside the scope of this thesis to assess their clinical utility or validity. A diagnostic test or prediction model would have clinical validity if it was demonstrated to work in a real-world setting and it would have clinical utility if patients whose treatment was informed by the test fared better on average than those whose treatment was not so informed (Kapur, Phillips, & Insel, 2012). While STRATA-G was not designed to assess clinical utility, partial clinical validity can be inferred because data was collected from patients undergoing routine clinical treatment. However, replication in independent samples collected from clinical settings is needed to draw any definitive conclusions about the clinical validity of these results. These studies also need to report their findings according to Standards for Reporting of Diagnostic Accuracy (STARD) statement (Bossuyt et al., 2003) so the prediction model is not used outside of the context in which it was validated (Streiner, 2003).

If the current model retained at least the same predictive performance in an independent sample it would be cost-effective to use clozapine as a second-line treatment for patients identified as TR by the model (Jin et al., 2019). Especially, given the low cost to healthcare organisations, and patients, in acquiring clinical and demographic data. Even including a polygenic risk score, the model would still

be cost effective as Jin et al. (2019) factored in the cost of a genetic predictor when modelling the health-economics associated with predicting TR. Polygenic risk scores are an ideal biological predictor as they require only a small blood sample, the cost of genotyping is constantly decreasing, and calculating the risk score is relatively simple (Lewis & Vassos, 2017). Furthermore, polygenic risk scores use DNA, which is stable from birth, and are therefore unconfounded by acute environmental factors (Lewis & Vassos, 2017). In terms of using a polygenic risk score in clinical practice, evidence suggests that genetic testing for psychiatric disorders and prognosis is seen favourably by patients, their families, and psychiatrists (Jenkins & Arribas-Ayllon, 2016; Lawrence & Appelbaum, 2011). These studies did however highlight a concern that, given the diagnostic instability in psychiatry, communication of prognostic risk lacks precision. This is not necessarily limited to a genetics-based predictive model and could be a concern in using any predictive model in a clinical setting. Particularly in psychosis-spectrum disorders where diagnoses are not mutually exclusive (Guloksuz & Van Os, 2018) and especially liable to change in first episode patients (Fusar-Poli et al., 2016).

Another concern is that individuals would interpret the results of a predictive model as deterministic. Personalised medicine – a unique intervention customised just for the given individual – may be unlikely to ever become a reality in psychiatry (Kapur et al., 2012). Instead Kapur et al. (2012) consider the possibility of stratified medicine where a biomarker or predictive model is identified that can stratify a broad-illness phenotype into a finite number of treatment-relevant subgroups. My prediction model is a form of the latter, and this distinction is important to consider when interpreting the results for individual patients. The model specifies the risk of being treatment resistant in the form of a probability and using a cut-off that splits patients into TR and NTR. This cut off was identifying during model-development and is similar to the prevalence of TR in the sample. The model is not able to determine whether an individual is TR with any diagnostic certainty. Rather, it identifies individual who are highly likely to be, or at risk of being, treatment resistant. It would be important to emphasize this if the model was used to predict TR at an individual, rather than group, level.

12.3.3. Understanding the aetiology of treatment resistance and novel drug development

The use of predictors or a prediction model is not limited to clinical practice. They can also be used in randomised controlled trials as inclusion criteria (Vickers, Kramer, & Baker, 2006) or to adjust for covariates (Hernández, Steyerberg, & Habbema, 2004). In recent years, pharmaceutical studies have developed drugs to target glutamate dysfunction in psychosis. Beck, Javitt, and Howes (2016) reviewed these studies and concluded that the majority have been unsuccessful: either trials have failed to meet their primary endpoints or are considered failed studies because patients in the treatment arm did not improve in comparison to patients randomised to placebo. Beck et al. (2016) discuss a number of reasons why these trials have been so unsuccessful, including suboptimal drug plasma levels and the strong placebo responses found in clinical trials. They argue that future clinical trials should focus on patients earlier in the course of their illness or a subgroup of patients who do not respond to dopamine-blocking antipsychotics. My prediction model would allow future clinical trials to recruit patients at first episode, or after only one antipsychotic trial, and to specifically recruit a subgroup at higher risk of being treatment resistant. Even before designing such a study, the prediction model could be applied to existing data from clinical trials of novel compounds to perform subgroup analysis.

Using prediction modelling to adjust for covariates should also apply to epidemiological or neuroimaging studies of treatment resistance. In a simulated randomised control trial (Hernández et al., 2004) found that adjusting for predictors increased the statistical power to detect a treatment effect without increasing the risk of type I error. If this was translated to neuroimaging studies, adjusting for predictive covariates could increase the statistical power of studies to detect neurochemical differences between TR and NTR patients. This could be one statistical solution to counterbalance the relatively small sample sizes prevalent in this type of research.

12.4. Strengths & Limitations

12.4.1. Longitudinal cohort studies

How the results of this thesis are interpreted is, in part, determined by whether the STRATA-G sample is representative of first episode psychosis patients and

treatment resistant patients in general. It is therefore important to discuss any biases inherent in the recruitment of first episode psychosis patients, as well as attrition bias, and whether the criteria for inclusion in STRATA-G has also limited the representativeness of the sample.

Many first episode psychosis studies use a tightly defined catchment area and time frame in which to recruit participants; this limits the impact of selection bias. However, not everyone identified as meeting the inclusion criteria will agree to take part in the study. For example, in the UK-based AESOP study, only 66% of the 592 patients identified as meeting inclusion criteria agreed to take part (Morgan et al., 2006). 10% either could not be contacted or did not speak English, while 13% simply declined to take part. Whether people who decline to take part in research are systematically different to those consent to take part is unknown. In examining participation in drug trials, lower symptom severity is associated with participation: Lester and Wilson (1999) found that positive symptoms increases suspiciousness towards researchers while negative symptoms decrease motivation to participate.

Another caveat in using longitudinal data is that the loss of participants at follow-up also biases the sample. Returning to the AESOP study as an example: follow-up was attempted for 532 patients in the baseline sample but 7% had died and 6% has emigrated (Revier et al., 2015). The researchers attempted to trace 458 participants, 10% could not be traced and of the 90% who were found, only 53% agreed to be re-interviewed (1% lacked capacity to consent and 46% could either not be contacted or declined to be re-interviewed). Previous work using the Istanbul data has found that psychosis patients are more likely to drop out if they are less educated, more suspicious, less compliant with medication, and have a later illness onset (Üçok et al., 2007).

An individual's disinclination to take part in research studies, and attrition, can not only bias the sample collected but also reduce the overall sample size and consequently limit statistical power. Researchers attempt to create a representative and well powered sample by increasing the number of participants they recruit at baseline and using exhaustive strategies to follow-up participants. The majority of

studies included in STRATA-G recruited participants from all the clinical services within large geographical areas. One exception is the study in Istanbul, where participants are recruited from only one hospital, but to compensate recruitment into this study has been ongoing since 1997.

For STRATA-G, I only included participants from the original cohort who had at least one year follow-up and where treatment resistance status could be determined. This was in part to comply with data protection legislation by only including data from the original research studies that was necessary to meet the aims of the STRATA-G project. Therefore, I have not been able compare participants included in STRATA-G to those recruited into the original studies; it is possible that, due to selection bias, STRATA-G is not fully representative of first episode psychosis cohorts. There is always a risk in analysing data collected by others, or collected for unrelated research questions, that you are not aware of all the sources of variability or bias in the data (Lambert & Black, 2012). Yet, despite these biases inherent in longitudinal cohort studies, they remain the best observational research design from which to identify risk factors of outcomes and perform prediction modelling (Chatterjee, Shi, & García-Closas, 2016). Furthermore, STRATA-G is the largest and most representative first episode psychosis cohort that has been used to investigate TR. By combining multiple studies from different countries, my findings are generalisable to a far wider range of patients with psychosis than if I had used only one cohort.

12.4.2. Defining treatment resistance

An important concern in all treatment resistance research is misclassification. Using longitudinal observational data to determine treatment resistance means that the participants in STRATA-G may have been misclassified. A lack of adherence to medication or subtherapeutic plasma levels (McCutcheon et al., 2015) could result in NTR patients being incorrectly classified as TR. Neither adherence nor plasma levels were collected consistently across the cohorts which make up the STRATA-G sample. Therefore, I could not perform sensitivity analyses in the subgroups that had good adherence or had received antipsychotics at a dose that corresponded to therapeutic plasma levels. This limitation is in part mitigated by the substantial length of follow-up (between 1 and 10 years). The long follow-up

also ensures that it is unlikely that TR patients are misclassified as NTR because as there is plenty of time for a diagnosis of TR to be established. Like the studies identified in my systematic review, I used pragmatic criteria to define treatment resistance in the hope that any predictors identified would be generalisable to real-world, clinical settings where adherence, compliance, or drug plasma levels influence treatment and response.

In addition, I used clozapine (*or* two antipsychotic failures) to define treatment resistance. Patients who are classified as TR using clozapine prescription are also highly likely to meet the two treatment failures criteria. However, not all patients who meet the two treatment failures criteria will be prescribed clozapine, as some will refuse clozapine treatment after considering the side effect profile and some clinicians are reluctant to prescribe clozapine for the same reason (Nielsen, Dahm, Lublin, & Taylor, 2010). It follows that studies which use clozapine prescription alone to define treatment resistance are often criticised for underestimating the prevalence of TR, but the proportion of TR patients missed is likely to be small. Freitas et al. (2019) examined symptom profiles in TR and NTR patients using a sample of 1429 participants. Of the patients who did not meet the two treatment failures criteria, none were prescribed clozapine. Of the patients who did meet the two treatment failures criteria, 92 (6.4% of the total sample) were also prescribed clozapine in previous year, while 409 (28.62%) were not prescribed clozapine. By using both criteria for to define TR I can be sure that the number of misclassified patients will be relatively small.

A final limitation is the varying rates of TR observed across cohorts included in the STRATA-G sample. This is an expected limitation in data from international research consortium (Budin-Ljøsne et al., 2014), which may reflect how the demographics of patients vary depending on which patients are seen by clinical services, followed up by services, and meet criteria for TR. Although this violates the assumption of homogeneity, this is not an assumption required for logistic regressions. In Chapter 11, I meta-analysed results where possible to overcome this limitation. In addition, it could be supposed that individuals are not completely independent from one another i.e. individuals within one cohort are more similar to one another than to an individual from a different cohort. I have attempted to

remove this limitation using various statistical techniques. In my causal model, I attempted to use multi-level modelling for missing data imputation. The data did not converge i.e. the multi-level missing data imputation model did not fit the data. Instead I imputed missing data, without multi-level modelling, and applied the multi-level modelling method to subsamples of participants and variables. There was no difference between the two methods in terms of how well missing data was imputed. To control for difference between cohorts, I included the cohorts in the explanatory model as dummy variables, leaving out one cohort to be used as a reference category. On the other hand, the heterogeneity in TR prevalence rates may be beneficial in explanatory modelling, as the magnitude of effects may be more robust (Bühlmann & van de Geer, 2018). For my prediction model, we used random forests to impute missing data, which does not have the option for multi-level imputation. I considered using leave-one-out (cohort) cross validation for the model development, but the volume of missing data prevented this. Instead I used repeated cross-fold validation (5 folds, 50 repeats). I could not include cohort as potential predictor, because this would have little value as a predictor in future data. Instead, I computed the area under the curve (AUC) in each cohort separately and found some variation, but no cohort had an AUC less than 0.6 and the maximum AUC was 0.76. Despite this heterogeneity, my model still was able to predict TR.

12.5. Conclusions

In conclusion, my data suggests that a model, including nine clinical and demographic variables, collected when a patient first presents to clinical services for psychosis, can be used to predict treatment resistant psychosis better than chance. In addition, the data suggests there is a genetic architecture that is specific to treatment resistance and, when used to calculate a polygenic risk score, explains a small proportion of the variance in whether an individual is treatment resistant or non-treatment resistant to antipsychotic medication. Finally, I have found that the most robust risk factor for treatment resistance is younger age of psychosis onset, and that age of onset is not caused by a higher genetic liability for treatment resistance. Overall, this suggests that, after further replication and validation, treatment resistance could be identified much earlier in the disease course, than is currently the case in routine clinical practice, and patients given faster access to clozapine.

REFERENCES

13. References

1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74. doi:10.1038/nature15393

http://www.nature.com/nature/journal/v526/n7571/abs/nature15393.html#supplementary-information

Agid, O., Arenovich, T., Sajeev, G., Zipursky, R. B., Kapur, S., Foussias, G., & Remington, G. (2011). An algorithm-based approach to first-episode schizophrenia: response rates over 3 prospective antipsychotic trials with a retrospective data analysis. *J Clin Psychiatry, 72*(11), 1439-1444.

Agid, O., Remington, G., Kapur, S., Arenovich, T., & Zipursky, R. B. (2007). Early use of clozapine for poorly responding first-episode psychosis. *Journal of clinical psychopharmacology, 27*(4), 369-373.

Ajnakina, O., Lally, J., Di Forti, M., Kolliakou, A., Gardner-Sood, P., Lopez-Morinigo, J., . . . Vassos, E. (2017). Patterns of illness and care over the 5 years following onset of psychosis in different ethnic groups; the GAP-5 study. *Social Psychiatry and Psychiatric Epidemiology, 52*(9), 1101-1111. doi:10.1007/s00127-017-1417-6

Alameda, L., Golay, P., Baumann, P. S., Progin, P., Mebdouhi, N., Elowe, J., . . . Conus, P. (2017). Mild Depressive Symptoms Mediate the Impact of Childhood Trauma on Long-Term Functional Outcome in Early Psychosis Patients. *Schizophr Bull, 43*(5), 1027-1035. doi:10.1093/schbul/sbw163

Albert, N., Bertelsen, M., Thorup, A., Petersen, L., Jeppesen, P., Le Quack, P., . . . Nordentoft, M. (2011). Predictors of recovery from psychosis Analyses of clinical and social factors associated with recovery among patients with first-episode psychosis after 5 years. *Schizophr Res, 125*(2-3), 257-266. doi:10.1016/j.schres.2010.10.013

Altman, D. G., & Bland, J. M. (2003). Interaction revisited: the difference between two estimates. *BMJ, 326*(7382), 219. doi:10.1136/bmj.326.7382.219

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders DSM-IV-TR fourth edition (text revision).

Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nature protocols, 5*(9), 1564-1573. doi:10.1038/nprot.2010.116

Anderson, V. M., McIlwain, M. E., Kydd, R. R., & Russell, B. R. (2015). Does cognitive impairment in treatment-resistant and ultra-treatment-resistant schizophrenia differ from that in treatment responders? *Psychiatry Research, 230*(3), 811-818.

Andreasen, N. C. (1982). Negative symptoms in schizophrenia: definition and reliability. *Arch Gen Psychiatry, 39*(7), 784-788.

Andreasen, N. C. (1983). *The Scale for the Assessment of Negative Symptoms (SANS)*. Iowa City, Iowa: The University of Iowa.

Andreasen, N. C. (1984). *The Scale for the Assessment of Positive Symptoms (SAPS)*. Iowa City, Iowa: The University of Iowa.

Andreasen, N. C. (1989). The Scale for the Assessment of Negative Symptoms (SANS): Conceptual and Theoretical Foundations. *British Journal of Psychiatry, 155*(S7), 49-52. doi:10.1192/S0007125000291496

Andrews, A., Knapp, M., McCrone, P., Parsonage, M., & Trachtenberg, M. (2012). *Effective Interventions in Schizophrenia: the Economic Case. A Report Prepared for the Schizophrenia Commission*. Retrieved from London:

Anttila, S., Illi, A., Kampman, O., Mattila, K. M., Lehtimäki, T., & Leinonen, E. (2005). Lack of association between two polymorphisms of brain-derived neurotrophic factor and response to typical neuroleptics. *Journal of Neural Transmission, 112*(7), 885-890. doi:10.1007/s00702-004-0233-9

Anttila, S., Kampman, O., Illi, A., Rontu, R., Lehtimäki, T., & Leinonen, E. (2007). Association between 5-HT2A, TPH1 and GNB3 genotypes and response to typical neuroleptics: a serotonergic approach. *BMC Psychiatry, 7*(1), 22. doi:10.1186/1471-244X-7-22

Arsalan, A., Iqbal, Z., Tariq, M., Ayonrinde, O., Vincent, J. B., & Ayub, M. (2019). Association of smoked cannabis with treatment resistance in schizophrenia. *Psychiatry Research, 278*, 242-247. doi:https://doi.org/10.1016/j.psychres.2019.06.023

Athanasiu, L., Mattingsdal, M., Kahler, A. K., Brown, A., Gustafsson, O., Agartz, I., . . . Andreassen, O. A. (2010). Gene variants associated with schizophrenia in a Norwegian genome-wide study are replicated in a large European cohort. *J Psychiatr Res, 44*(12), 748-753. doi:10.1016/j.jpsychires.2010.02.002

Ayesa-Arriola, R., Teran, J. M. P., Morinigo, J. D. L., Rivero, M. C., Setien-Suero, E., Al-Halabi, S., . . . Crespo-Facorro, B. (2018). The dynamic relationship between insight and suicidal behavior in first episode psychosis patients over 3-year follow-up. *Eur Neuropsychopharmacol*. doi:10.1016/j.euroneuro.2018.05.005

Ayres, A. M., Busatto, G. F., Menezes, P. R., Schaufelberger, M. S., Coutinho, L., Murray, R. M., . . . Scazufca, M. (2007). Cognitive deficits in first-episode psychosis: a population-based study in Sao Paulo, Brazil. *Schizophrenia Research, 90*(1-3), 338-343.

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple Imputation by Chained Equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research, 20*(1), 40-49. doi:10.1002/mpr.329

Ballageer, T., Malla, A., Manchanda, R., Takhar, J., & Haricharan, R. (2005). Is adolescent-onset first-episode psychosis different from adult onset? *J Am Acad Child Adolesc Psychiatry, 44*(8), 782-789. doi:10.1097/01.chi.0000164591.55942.ea

Barrantes-Vidal, N., Grant, P., & Kwapil, T. R. (2015). The Role of Schizotypy in the Study of the Etiology of Schizophrenia Spectrum Disorders. *Schizophr Bull, 41*(suppl_2), S408-S416. doi:10.1093/schbul/sbu191

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual differences, 42*(5), 815-824.

Baumann, P. S., Crespi, S., Marion-Veyron, R., Solida, A., Thonney, J., Favrod, J., . . . Conus, P. (2013). Treatment and Early Intervention in Psychosis Program (TIPP-Lausanne): implementation of an early intervention programme for psychosis in Switzerland. *Early Intervention in Psychiatry, 7*(3), 322-328. doi:doi:10.1111/eip.12037

Beck, K., Javitt, D. C., & Howes, O. D. (2016). Targeting glutamate to treat schizophrenia: lessons from recent clinical studies. *Psychopharmacology, 233*(13), 2425-2428. doi:10.1007/s00213-016-4318-6

Beck, K., McCutcheon, R., Stephenson, L., Schilderman, M., Patel, N., Ramsay, R., & Howes, O. D. (2019). Prevalence of treatment-resistant psychoses in the community: A naturalistic study. *J Psychopharmacol*, 269881119855995. doi:10.1177/0269881119855995

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the royal statistical society. Series B (Methodological), 57*(1), 289-300. Retrieved from http://www.jstor.org/stable/2346101

Bergen, S. E., O'Dushlaine, C. T., Ripke, S., Lee, P. H., Ruderfer, D. M., Akterin, S., . . . Sullivan, P. F. (2012). Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol Psychiatry, 17*(9), 880-886. doi:10.1038/mp.2012.73

Berisa, T., & Pickrell, J. K. (2016). Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics (Oxford, England), 32*(2), 283-285. doi:10.1093/bioinformatics/btv546

Bilic, P., Jukic, V., Vilibic, M., Savic, A., & Bozina, N. (2014). Treatment-resistant schizophrenia and DAT and SERT polymorphisms. *Gene, 543*(1), 125-132. doi:https://doi.org/10.1016/j.gene.2014.03.050

Bobes, J., Garcia-Portilla, M. P., Bascaran, M. T., Saiz, P. A., & Bousoño, M. (2007). Quality of life in schizophrenic patients. *Dialogues in clinical neuroscience, 9*(2), 215-226. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17726919
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3181847/

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., . . . de Vet, H. C. W. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ, 326*(7379), 41-44. doi:10.1136/bmj.326.7379.41

Bozzatello, P., Bellino, S., & Rocca, P. (2019). Predictive factors of treatment resistance in first episode of psychosis: a systematic review. *Frontiers in Psychiatry, 10*, 67.

Brain, C., Kymes, S., DiBenedetti, D. B., Brevig, T., & Velligan, D. I. (2018). Experiences, attitudes, and perceptions of caregivers of individuals with treatment-resistant schizophrenia: a qualitative study. *BMC Psychiatry, 18*(1), 253. doi:10.1186/s12888-018-1833-5

Budin-Ljøsne, I., Isaeva, J., Maria Knoppers, B., Marie Tassé, A., Shen, H.-y., McCarthy, M. I., . . . Consortium, E. (2014). Data sharing in large research consortia: experiences and recommendations from ENGAGE. *European Journal of Human Genetics, 22*(3), 317-321. doi:10.1038/ejhg.2013.131

Bühlmann, P., & van de Geer, S. (2018). Statistics for big data: A perspective. *Statistics & Probability Letters, 136*, 37-41. doi:https://doi.org/10.1016/j.spl.2018.02.016

Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., . . . Neale, B. M. (2015a). An atlas of genetic correlations across human diseases and traits. *Nat Genet, 47*(11), 1236-1241. doi:10.1038/ng.3406

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics, C., . . . Neale, B. M. (2015b). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet, 47*(3), 291-295. doi:10.1038/ng.3211

http://www.nature.com/ng/journal/v47/n3/abs/ng.3211.html#supplementary-information

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., . . . Sollis, E. (2018). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic acids research, 47*(D1), D1005-D1012.

Bzdok, D., & Meyer-Lindenberg, A. (2018). Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(3), 223-230. doi:https://doi.org/10.1016/j.bpsc.2017.11.007

Carbon, M., & Correll, C. U. (2014). Clinical predictors of therapeutic response to antipsychotics in schizophrenia. *Dialogues in clinical neuroscience, 16*(4), 505-524. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4336920/

Carpenter Jr, W. T. (2007). Schizophrenia: Disease, Syndrome, or Dimensions? *Family Process, 46*(2), 199-206. doi:10.1111/j.1545-5300.2007.00204.x

Carter, L., Read, J., Pyle, M., Law, H., Emsley, R., & Morrison, A. (2018). Causal beliefs in people experiencing psychosis: The relationship to treatment accessed and the perceived helpfulness of treatment. *Psychology and Psychotherapy: Theory, Research and Practice, 91*(3), 332-344. doi:10.1111/papt.12163

Casey, P., & Corvin, A. (2008). The Clinical Impact of Substance Use in Schizophrenia: A Study in an Irish Population. *TSMJ, 9*, 14-17.

Castle, D., Jablensky, A., McGrath, J., Carr, V., Morgan, V., Waterreus, A., . . . Farmer, A. (2006). The diagnostic interview for psychoses (DIP): development, reliability and applications. *Psychological Medicine, 36*(1), 69-80.

Chan, S. K. W., Chang, W. C., Lee, E. H. M., Wong, G. H. Y., Hui, C. L. M., Chen, E. Y. H., & Sham, P. C. (2014). Predictors of treatment resistant schizophrenia-spectrum disorder: 10-year retrospective study of first-episode psychosis. *Early Intervention in Psychiatry, 8*, 78. doi:https://rfs1.fhb.gov.hk/app/fundedsearch/projectdetail.xhtml?id=1363

Chatterjee, N., Shi, J., & García-Closas, M. (2016). Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nature Reviews Genetics, 17*, 392. doi:10.1038/nrg.2016.27

https://www.nature.com/articles/nrg.2016.27#supplementary-information

Chen, Y.-L., Chen, K.-P., Chiu, C.-C., Tai, M.-H., & Lung, F.-W. (2018). Early predictors of poor treatment response in patients with schizophrenia treated with atypical antipsychotics. *BMC Psychiatry, 18*(1), 376. doi:10.1186/s12888-018-1950-1

Chiliza, B., Asmal, L., Kilian, S., Phahladira, L., & Emsley, R. (2015). Rate and predictors of non-response to first-line antipsychotic treatment in first-episode schizophrenia. *30*(3), 173-182. doi:10.1002/hup.2469

Choong, E., Quteineh, L., Cardinaux, J. R., Gholam-Rezaee, M., Vandenberghe, F., Dobrinas, M., . . . Eap, C. B. (2013). Influence of CRTC1 polymorphisms on body mass index and fat mass in psychiatric patients and the general adult population. *JAMA Psychiatry, 70*(10), 1011-1019. doi:10.1001/jamapsychiatry.2013.187

Choong, E., Solida, A., Lechaire, C., Conus, P., & Eap, C. B. (2008). Follow-up of the metabolic syndrome induced by atypical antipsychotics:

recommendations and pharmacogenetics perspectives. *Revue médicale suisse, 4*(171), 1994-1996,1998-1999. Retrieved from https://archive-ouverte.unige.ch/unige:16944

Ciarleglio, A. J., Brucato, G., Masucci, M. D., Altschuler, R., Colibazzi, T., Corcoran, C. M., . . . Girgis, R. R. (2018). A predictive model for conversion to psychosis in clinical high-risk patients. *Psychol Med*, 1-10. doi:10.1017/s003329171800171x

Clarke, G. M., Anderson, C. A., Pettersson, F. H., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2011). Basic statistical analysis in genetic case-control studies. *Nature Protocols, 6*(2), 121-133. doi:doi:10.1038/nprot.2010.182

Cleynen, I., Boucher, G., Jostins, L., Schumm, L. P., Zeissig, S., Ahmad, T., . . . Lees, C. W. (2016). Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *The Lancet, 387*(10014), 156-167. doi:https://doi.org/10.1016/S0140-6736(15)00465-1

Coleman, J. R. I., Euesden, J., Patel, H., Folarin, A. A., Newhouse, S., & Breen, G. (2016). Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Briefings in functional genomics, 15*(4), 298-304. doi:10.1093/bfgp/elv037

Conley, R. R., & Kelly, D. L. (2001). Management of treatment resistance in schizophrenia. *Biol Psychiatry, 50*(11), 898-911. doi:https://doi.org/10.1016/S0006-3223(01)01271-9

Conomos, M. P., Miller, M. B., & Thornton, T. A. (2015). Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. *Genet Epidemiol, 39*(4), 276-293. doi:10.1002/gepi.21896

Conomos, M. P., Reiner, A. P., Weir, B. S., & Thornton, T. A. (2016). Model-free Estimation of Recent Genetic Relatedness. *American journal of human genetics, 98*(1), 127-148. doi:10.1016/j.ajhg.2015.11.022

Crespo-Facorro, B., Pelayo-Teran, J. M., Perez-Iglesias, R., Ramirez-Bonilla, M., Martinez-Garcia, O., Pardo-Garcia, G., & Vazquez-Barquero, J. L. (2007). Predictors of acute treatment response in patients with a first episode of non-affective psychosis: sociodemographics, premorbid and clinical variables. *J Psychiatr Res, 41*(8), 659-666. doi:10.1016/j.jpsychires.2006.05.002

Cunill, R., Castells, X., & Simeon, D. (2009). Relationships between obsessive-compulsive symptomatology and severity of psychosis in schizophrenia: a systematic review and meta-analysis. In.

Das, S., Forer, L., Schonherr, S., Sidore, C., Locke, A. E., Kwong, A., . . . Fuchsberger, C. (2016). Next-generation genotype imputation service and methods. *Nat Genet, 48*(10), 1284-1287. doi:10.1038/ng.3656

Dazzan, P., Morgan, K. D., Orr, K., Hutchinson, G., Chitnis, X., Suckling, J., . . . Murray, R. M. (2005). Different Effects of Typical and Atypical Antipsychotics on Grey Matter in First Episode Psychosis: the ÆSOP Study. *Neuropsychopharmacology, 30*, 765. doi:10.1038/sj.npp.1300603

de Bartolomeis, A., Balletta, R., Giordano, S., Buonaguro, E. F., Latte, G., & Iasevoli, F. (2013). Differential cognitive performances between schizophrenic responders and non-responders to antipsychotics: Correlation with course of the illness, psychopathology, attitude to the treatment and antipsychotics doses. *Psychiatry Research, 210*(2), 387-395. doi:https://doi.org/10.1016/j.psychres.2013.06.042

Dean, K., Fearon, P., Morgan, K., Hutchinson, G., Orr, K., Chitnis, X., . . . Dazzan, P. (2018). Grey matter correlates of minor physical anomalies in the ÆSOP first-episode psychosis study. *British Journal of Psychiatry, 189*(3), 221-228. doi:10.1192/bjp.bp.105.016337

Dean, K., & Murray, R. M. (2005). Environmental risk factors for psychosis. *Dialogues in clinical neuroscience, 7*(1), 69-80. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16060597
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3181718/

Delacretaz, A., Preisig, M., Vandenberghe, F., Saigi Morgui, N., Quteineh, L., Choong, E., . . . Eap, C. B. (2015). Influence of MCHR2 and MCHR2-AS1 Genetic Polymorphisms on Body Mass Index in Psychiatric Patients and In Population-Based Subjects with Present or Past Atypical Depression. *PLoS One, 10*(10), e0139155. doi:10.1371/journal.pone.0139155

Delaneau, O., Zagury, J.-F., & Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nature Methods, 10*(1), 5-6.

Demjaha, A., Egerton, A., Murray, R. M., Kapur, S., Howes, O. D., Stone, J. M., & McGuire, P. K. (2014). Antipsychotic treatment resistance in schizophrenia associated with elevated glutamate levels but normal dopamine function. *Biol Psychiatry, 75*(5), e11-e13.

Demjaha, A., Lappin, J. M., Stahl, D., Patel, M. X., MacCabe, J. H., Howes, O. D., . . . Murray, R. M. (2017). Antipsychotic treatment resistance in first-episode psychosis: prevalence, subtypes and predictors. *Psychological Medicine, 47*(11), 1981-1989. doi:10.1017/S0033291717000435

Di Forti, M., Marconi, A., Carra, E., Fraietta, S., Trotta, A., Bonomo, M., . . . Murray, R. M. (2015). Proportion of patients in south London with first-episode psychosis attributable to use of high potency cannabis: a case-control study. *The Lancet Psychiatry, 2*(3), 233-238. doi:10.1016/S2215-0366(14)00117-5

Di Forti, M., Morgan, C., Dazzan, P., Pariante, C., Mondelli, V., Marques, T. R., . . . Murray, R. M. (2009). High-potency cannabis and the risk of psychosis. *The British Journal of Psychiatry, 195*(6), 488-491. doi:10.1192/bjp.bp.109.064220

Downs, J., Dean, H., Lechler, S., Sears, N., Patel, R., Shetty, H., . . . Pina-Camacho, L. (2019). Negative Symptoms in Early-Onset Psychosis and Their Association With Antipsychotic Treatment Failure. *Schizophr Bull, 45*(1), 69-79. doi:10.1093/schbul/sbx197

Downs, J. M., Lechler, S., Dean, H., Sears, N., Patel, R., Shetty, H., . . . Pina-Camacho, L. (2017). The Association Between Comorbid Autism Spectrum Disorders and Antipsychotic Treatment Failure in Early-Onset Psychosis: A Historical Cohort Study Using Electronic Health Records. *J Clin Psychiatry, 78*(9), e1233-e1241. doi:10.4088/JCP.16m11422

Drake, R. E., Osher, F. C., Noordsy, D. L., Hurlbut, S. C., Teague, G. B., & Beaudett, M. S. (1990). Diagnosis of Alcohol Use Disorders in Schizophrenia. *Schizophr Bull, 16*(1), 57-67. doi:10.1093/schbul/16.1.57

Dubey, R., Zhou, J., Wang, Y., Thompson, P. M., & Ye, J. (2014). Analysis of sampling techniques for imbalanced data: An n = 648 ADNI study. *Neuroimage, 87*, 220-241. doi:10.1016/j.neuroimage.2013.10.005

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genet, 9*(3), e1003348. doi:10.1371/journal.pgen.1003348

Egerton, A., Brugger, S., Raffin, M., Barker, G. J., Lythgoe, D. J., McGuire, P. K., & Stone, J. M. (2012). Anterior cingulate glutamate levels related to clinical status following treatment in first-episode schizophrenia. *Neuropsychopharmacology, 37*(11), 2515.

Egerton, A., Murphy, A., Donocik, J., Anton, A., Barker, G. J., Collier, T., . . . Howes, O. D. (Unpublished). Dopamine and Glutamate in Antipsychotic-Responsive compared to
Antipsychotic Non-Responsive Psychosis: A Multicentre 18F-DOPA PET
and 1H-MRS Study (STRATA).

Ekholm, B., Ekholm, A., Adolfsson, R., Vares, M., Ösby, U., Sedvall, G. C., & Jönsson, E. G. (2005). Evaluation of diagnostic procedures in Swedish patients with schizophrenia and related psychoses. *Nordic Journal of Psychiatry, 59*(6), 457-464. doi:10.1080/08039480500360906

Elkis, H., & Buckley, P. F. (2016). Treatment-Resistant Schizophrenia. *Psychiatric Clinics, 39*(2), 239-265. doi:10.1016/j.psc.2016.01.006

Emsley, R., Chiliza, B., Asmal, L., & Harvey, B. H. (2013). The nature of relapse in schizophrenia. *BMC Psychiatry, 13*, 50-50. doi:10.1186/1471-244X-13-50

Emsley, R., Oosthuizen, P. P., Kidd, M., Koen, L., Niehaus, D. J., & Turner, H. J. (2006). Remission in first-episode psychosis: predictor variables and symptom improvement patterns. *J Clin Psychiatry, 67*(11), 1707-1712.

Eranti, S. V., MacCabe, J. H., Bundy, H., & Murray, R. M. (2013). Gender difference in age at onset of schizophrenia: a meta-analysis. *Psychol Med, 43*(1), 155-167. doi:10.1017/s003329171200089x

Esterberg, M. L., Trotman, H. D., Holtzman, C., Compton, M. T., & Walker, E. F. (2010). The impact of a family history of psychosis on age-at-onset and positive and negative symptoms of schizophrenia: A meta-analysis. *Schizophrenia Research, 120*(1), 121-130. doi:https://doi.org/10.1016/j.schres.2010.01.011

Euesden, J., Lewis, C. M., & O'Reilly, P. F. (2014). PRSice: polygenic risk score software. *Bioinformatics (Oxford, England), 31*(9), 1466-1468.

Faerden, A., Barrett, E. A., Nesvåg, R., Friis, S., Finset, A., Marder, S. R., . . . Melle, I. (2013). Apathy, poor verbal memory and male gender predict lower psychosocial functioning one year after the first treatment of psychosis. *Psychiatry Research, 210*(1), 55-61. doi:https://doi.org/10.1016/j.psychres.2013.02.007

Faerden, A., Nesvag, R., Barrett, E. A., Agartz, I., Finset, A., Friis, S., . . . Melle, I. (2008). Assessing apathy: the use of the Apathy Evaluation Scale in first episode psychosis. *Eur Psychiatry, 23*(1), 33-39. doi:10.1016/j.eurpsy.2007.09.002

Fearon, P., Kirkbride, J. B., Morgan, C., Dazzan, P., Morgan, K., Lloyd, T., . . . Murray, R. M. (2006). Incidence of schizophrenia and other psychoses in ethnic minority groups: results from the MRC AESOP Study. *Psychol Med, 36*(11), 1541-1550. doi:10.1017/s0033291706008774

Frank, J., Lang, M., Witt, S., Strohmaier, J., Rujescu, D., Cichon, S., . . . Ripke, S. (2015a). Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients. *Mol Psychiatry, 20*, 150–151. doi:10.1038/mp.2014.56

Frank, J., Lang, M., Witt, S. H., Strohmaier, J., Rujescu, D., Cichon, S., . . . Rietschel, M. (2015b). Identification of increased genetic risk scores for

schizophrenia in treatment-resistant patients. *Mol Psychiatry, 20*(2), 150-151. doi:10.1038/mp.2014.56

Freitas, R., Dos Santos, B., Altamura, C., Bernasconi, C., Corral, R., Evans, J., . . . Elkis, H. (2019). Can the Positive and Negative Syndrome scale (PANSS) differentiate treatment-resistant from non-treatment-resistant schizophrenia? A factor analytic investigation based on data from the Pattern cohort study. *Psychiatry Res, 276*, 210-217. doi:10.1016/j.psychres.2019.05.002

Frydecka, D., Beszłej, J. A., Gościmski, P., Kiejna, A., & Misiak, B. (2016). Profiling cognitive impairment in treatment-resistant schizophrenia patients. *Psychiatry Research, 235*, 133-138. doi:https://doi.org/10.1016/j.psychres.2015.11.028

Fusar-Poli, P., Cappucciati, M., Rutigliano, G., Heslin, M., Stahl, D., Brittenden, Z., . . . Carpenter, W. T. (2016). Diagnostic Stability of ICD/DSM First Episode Psychosis Diagnoses: Meta-analysis. *Schizophr Bull, 42*(6), 1395-1406. doi:10.1093/schbul/sbw020

Fusar-Poli, P., Papanastasiou, E., Stahl, D., Rocchetti, M., Carpenter, W., Shergill, S., & McGuire, P. (2015). Treatments of Negative Symptoms in Schizophrenia: Meta-Analysis of 168 Randomized Placebo-Controlled Trials. *Schizophr Bull, 41*(4), 892-899. doi:10.1093/schbul/sbu170

Giambartolomei, C., Vukcevic, D., Schadt, E. E., Franke, L., Hingorani, A. D., Wallace, C., & Plagnol, V. (2014). Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet, 10*(5), e1004383. doi:10.1371/journal.pgen.1004383

Gillespie, A. L., Samanaite, R., Mill, J., Egerton, A., & MacCabe, J. H. (2017). Is treatment-resistant schizophrenia categorically distinct from treatment-responsive schizophrenia? a systematic review. *BMC Psychiatry, 17*(1), 12. doi:10.1186/s12888-016-1177-y

Goes, F. S., McGrath, J., Avramopoulos, D., Wolyniec, P., Pirooznia, M., Ruczinski, I., . . . Peters, I. (2015). Genome-wide association study of schizophrenia in Ashkenazi Jews. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 168*(8), 649-659.

Golay, P., Alameda, L., Baumann, P., Elowe, J., Progin, P., Polari, A., & Conus, P. (2016). Duration of untreated psychosis: Impact of the definition of treatment onset on its predictive value over three years of treatment. *J Psychiatr Res, 77*, 15-21. doi:10.1016/j.jpsychires.2016.02.017

Goldberger, C., Gourion, D., Leroy, S., Schürhoff, F., Bourdel, M.-C., Leboyer, M., & Krebs, M.-O. (2005). Population-based and family-based association study of 5′UTR polymorphism of the reelin gene and schizophrenia. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 137B*(1), 51-55. doi:10.1002/ajmg.b.30191

Goldstein, M. E., Anderson, V. M., Pillai, A., Kydd, R. R., & Russell, B. R. (2015). Glutamatergic neurometabolites in clozapine-responsive and-resistant schizophrenia. *International Journal of Neuropsychopharmacology, 18*(6).

Gottesman, I. I., & Shields, J. (1972). A polygenic theory of schizophrenia. *International Journal of Mental Health, 1*(1-2), 107-115.

Guloksuz, S., & Van Os, J. (2018). The slow death of the concept of schizophrenia and the painful birth of the psychosis spectrum. *Psychological Medicine, 48*(2), 229.

Gusev, A., Ko, A., Shi, H., Bhatia, G., Chung, W., Penninx, B. W. J. H., . . . Pasaniuc, B. (2016). Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet, 48*(3), 245-252. doi:10.1038/ng.3506

Gutierrez-Galve, L., Chu, E. M., Leeson, V. C., Price, G., Barnes, T. R., Joyce, E. M., & Ron, M. A. (2015). A longitudinal study of cortical changes and their cognitive correlates in patients followed up after first-episode psychosis. *Psychol Med, 45*(1), 205-216. doi:10.1017/s0033291714001433

Gutierrez-Galve, L., Wheeler-Kingshott, C. A., Altmann, D. R., Price, G., Chu, E. M., Leeson, V. C., . . . Ron, M. A. (2010). Changes in the frontotemporal cortex and cognitive correlates in first-episode psychosis. *Biol Psychiatry, 68*(1), 51-60. doi:10.1016/j.biopsych.2010.03.019

Häfner, H. (2014). The concept of schizophrenia. *Advances in Psychiatry, 2014*. doi:10.1155/2014/929434

Hamshere, M. L., Walters, J. T. R., Smith, R., Richards, A. L., Green, E., Grozeva, D., . . . O'Donovan, M. C. (2013). Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Mol Psychiatry, 18*(6), 708-712. doi:10.1038/mp.2012.67

Hare, E., Glahn, D. C., Dassori, A., Raventos, H., Nicolini, H., Ontiveros, A., . . . Muñoz, R. (2010). Heritability of age of onset of psychosis in schizophrenia. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 153*(1), 298-302.

Harrell, F. E., Jr., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med, 15*(4), 361-387. doi:10.1002/(sici)1097-0258(19960229)15:4<361::aid-sim168>3.0.co;2-4

Harvey, R. C., James, A. C., & Shields, G. E. (2016). A Systematic Review and Network Meta-Analysis to Assess the Relative Efficacy of Antipsychotics for the Treatment of Positive and Negative Symptoms in Early-Onset Schizophrenia. *CNS Drugs, 30*(1), 27-39. doi:10.1007/s40263-015-0308-1

Hassan, A. N., & De Luca, V. (2015). The effect of lifetime adversities on resistance to antipsychotic treatment in schizophrenia patients. *Schizophrenia Research, 161*(2-3), 496-500.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. New York, NY: Springer.

Hayduk, L., Cummings, G., Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! testing! one, two, three–Testing the theory in structural equation models! *Personality and Individual differences, 42*(5), 841-850.

Hemingway, H., Croft, P., Perel, P., Hayden, J. A., Abrams, K., Timmis, A., . . . Riley, R. D. (2013). Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *346*, e5595. doi:10.1136/bmj.e5595 %J BMJ : British Medical Journal

Hernández, A. V., Steyerberg, E. W., & Habbema, J. D. F. (2004). Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements. *Journal of Clinical Epidemiology, 57*(5), 454-460.

Hickling, L. M., Ortiz-García de la Foz, V., Ayesa-Arriola, R., Crespo-Facorro, B., McGuire, P., & Perez-Iglesias, R. (2017). The effects of tobacco smoking on age of onset of psychosis and psychotic symptoms in a first-episode psychosis population. *Addiction, 112*(3), 526-532.

Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge, T. M., . . . Glenthoj, B. (2017). Is an Early Age at Illness Onset in Schizophrenia Associated With Increased Genetic Susceptibility? Analysis of Data From the Nationwide Danish Twin Register. *EBioMedicine, 18*, 320-326. doi:10.1016/j.ebiom.2017.04.002

Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge, T. M., . . . Glenthøj, B. (2018). Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biol Psychiatry, 83*(6), 492-498. doi:https://doi.org/10.1016/j.biopsych.2017.08.017

Horsdal, H. T., Wimberley, T., Benros, M. E., & Gasse, C. (2017a). C-reactive protein levels and treatment resistance in schizophrenia—A Danish population-based cohort study. *Human Psychopharmacology: Clinical and Experimental, 32*(6), e2632-n/a. doi:10.1002/hup.2632

Horsdal, H. T., Wimberley, T., Kohler-Forsberg, O., Baandrup, L., & Gasse, C. (2018). Association between global functioning at first schizophrenia diagnosis and treatment resistance. *Early Interv Psychiatry, 12*(6), 1198-1202. doi:10.1111/eip.12522

Horsdal, H. T., Wimberley, T., Köhler-Forsberg, O., Baandrup, L., & Gasse, C. (2017b). Association between global functioning at first schizophrenia diagnosis and treatment resistance. *Early Intervention in Psychiatry*, n/a-n/a. doi:10.1111/eip.12522

Hotta, Y., Ohnuma, T., Hanzawa, R., Shibata, N., Maeshima, H., Baba, H., . . . Arai, H. (2011). Association study between Disrupted-in-Schizophrenia-1 (DISC1) and Japanese patients with treatment-resistant schizophrenia (TRS). *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(2), 636-639. doi:https://doi.org/10.1016/j.pnpbp.2011.01.011

Howes, O. D., & Kapur, S. (2009). The Dopamine Hypothesis of Schizophrenia: Version III—The Final Common Pathway. *Schizophr Bull, 35*(3), 549-562. doi:10.1093/schbul/sbp006

Howes, O. D., McCutcheon, R., Agid, O., de Bartolomeis, A., van Beveren, N. J., Birnbaum, M. L., . . . Correll, C. U. (2017). Treatment-resistant schizophrenia: Treatment Response and Resistance in Psychosis (TRRIP) working group consensus guidelines on diagnosis and terminology. *The American Journal of Psychiatry, 174*(3), 216-229. doi:http://dx.doi.org/10.1176/appi.ajp.2016.16050503

Howes, O. D., & Murray, R. M. (2014). Schizophrenia: an integrated sociodevelopmental-cognitive model. *Lancet (London, England), 383*(9929), 1677-1687. doi:10.1016/S0140-6736(13)62036-X

Howes, O. D., Vergunst, F., Gee, S., McGuire, P., Kapur, S., & Taylor, D. (2012). Adherence to treatment guidelines in clinical practice: study of antipsychotic treatment prior to clozapine initiation. *British Journal of Psychiatry, 201*(6), 481-485. doi:10.1192/bjp.bp.111.105833

Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., & Abecasis, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet, 44*(8), 955-959.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6*(1), 1-55.

Huddy, V. C., Clark, L., Harrison, I., Ron, M. A., Moutoussis, M., Barnes, T. R., & Joyce, E. M. (2013). Reflection impulsivity and response inhibition in first-episode psychosis: relationship to cannabis use. *Psychol Med, 43*(10), 2097-2107. doi:10.1017/s0033291712003054

Huddy, V. C., Hodgson, T. L., Kapasi, M., Mutsatsa, S. H., Harrison, I., Barnes, T. R. E., & Joyce, E. M. (2007). Gaze strategies during planning in first-episode psychosis. *J Abnorm Psychol, 116*(3), 589-598. doi:10.1037/0021-843x.116.3.589

Hui, C. L., Honer, W. G., Lee, E. H., Chang, W. C., Chan, S. K., Chen, E. S., . . . Yeung, W. S. (2018). Long-term effects of discontinuation from antipsychotic maintenance following first-episode schizophrenia and related disorders: a 10 year follow-up of a randomised, double-blind trial. *The Lancet Psychiatry, 5*(5), 432-442.

Iasevoli, F., Avagliano, C., Altavilla, B., Barone, A., Ciccarelli, M., D'Ambrosio, L., . . . de Bartolomeis, A. (2018a). Evaluation of a few discrete clinical markers may predict categorization of actively symptomatic non-acute schizophrenia patients as treatment resistant or responders: A study by ROC curve analysis and multivariate analyses. *Psychiatry Res, 269*, 481-493. doi:10.1016/j.psychres.2018.08.109

Iasevoli, F., Avagliano, C., Altavilla, B., Barone, A., D'Ambrosio, L., Matrone, M., . . . de Bartolomeis, A. (2018b). Disease Severity in Treatment Resistant Schizophrenia Patients Is Mainly Affected by Negative Symptoms, Which Mediate the Effects of Cognitive Dysfunctions and Neurological Soft Signs. *Frontiers in Psychiatry, 9*(553). doi:10.3389/fpsyt.2018.00553

Iasevoli, F., Balletta, R., Gilardi, V., Giordano, S., & De Bartolomeis, A. (2013). Tobacco smoking in treatment-resistant schizophrenia patients is associated with impaired cognitive functioning, more severe negative symptoms, and poorer social adjustment. *Neuropsychiatric disease and treatment, 9*, 1113.

Iasevoli, F., D'Ambrosio, L., Notar Francesco, D., Razzino, E., Buonaguro, E. F., Giordano, S., . . . de Bartolomeis, A. (2018c). Clinical evaluation of functional capacity in treatment resistant schizophrenia patients: Comparison and differences with non-resistant schizophrenia patients. *Schizophrenia Research, 202*, 217-225. doi:https://doi.org/10.1016/j.schres.2018.06.030

Iasevoli, F., Giordano, S., Balletta, R., Latte, G., Formato, M. V., Prinzivalli, E., . . . de Bartolomeis, A. (2016). Treatment resistant schizophrenia is associated with the worst community functioning among severely-ill highly-disabling psychiatric conditions and is the most relevant predictor of poorer achievements in functional milestones. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 65*, 34-48.

Ikeda, M., Takahashi, A., Kamatani, Y., Momozawa, Y., Saito, T., Kondo, K., . . . Iwata, N. (2019). Genome-Wide Association Study Detected Novel Susceptibility Genes for Schizophrenia and Shared Trans-Populations/Diseases Genetic Effect. *Schizophr Bull, 45*(4), 824-834. doi:10.1093/schbul/sby140

Immonen, J., Jääskeläinen, E., Korpela, H., & Miettunen, J. (2017). Age at onset and the outcomes of schizophrenia: A systematic review and meta-analysis. *Early Intervention in Psychiatry, 11*(6), 453-460. doi:10.1111/eip.12412

Inada, T., Nakamura, A., & Iijima, Y. (2003). Relationship between catechol-O-methyltransferase polymorphism and treatment-resistant schizophrenia. *American Journal of Medical Genetics - Neuropsychiatric Genetics, 120 B*(1), 35-39. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-0042322722&partnerID=40&md5=76e4b3c4039e00243c9d9ac5acad79bb

International HapMap 3 Consortium. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature, 467*(7311), 52-58.

Jakobsen, K. D., Frederiksen, J. N., Hansen, T., Jansson, L. B., Parnas, J., & Werge, T. (2005). Reliability of clinical ICD-10 schizophrenia diagnoses. *Nordic Journal of Psychiatry, 59*(3), 209-212. doi:10.1080/08039480510027698

Jenkins, S., & Arribas-Ayllon, M. (2016). Genetic counselling for psychiatric disorders: accounts of psychiatric health professionals in the United Kingdom. *Journal of genetic counseling, 25*(6), 1243-1255.

Jia, P., Jayathilake, K., Zhao, Z., & Meltzer, H. Y. (2011). Association of FAS, a TNF-α receptor gene, with treatment resistant schizophrenia. *Schizophrenia Research, 129*(2), 211-212. doi:https://doi.org/10.1016/j.schres.2011.04.013

Jin, H., McCrone, P., & MacCabe, J. H. (2019). Stratified medicine in schizophrenia: how accurate would a test of drug response need to be to achieve cost-effective improvements in quality of life? *Eur J Health Econ*. doi:10.1007/s10198-019-01108-4

Jones, I., Scourfield, J., McCandless, F., & Craddock, N. (2002). Attitudes towards future testing for bipolar disorder susceptibility genes: a preliminary investigation. *Journal of affective disorders, 71*(1), 189-193. doi:https://doi.org/10.1016/S0165-0327(00)00384-0

Jongsma, H. E., Gayer-Anderson, C., Lasalvia, A., Quattrone, D., Mule, A., Szoke, A., . . . Kirkbride, J. B. (2018). Treated Incidence of Psychotic Disorders in the Multinational EU-GEI Study. *JAMA Psychiatry, 75*(1), 36-46. doi:10.1001/jamapsychiatry.2017.3554

Joober, R., Benkelfat, C., Brisebois, K., Toulouse, A., Turecki, G., Lal, S., . . . Rouleau, G. A. (1999). T102C polymorphism in the 5HT2A gene and schizophrenia: relation to phenotype and drug response variability. *Journal of psychiatry & neuroscience : JPN, 24*(2), 141-146. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10212557
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1188995/

Joober, R., Rouleau, G. A., Lal, S., Bloom, D., Lalonde, P., Labelle, A., & Benkelfat, C. (2005). Increased prevalence of schizophrenia spectrum disorders in relatives of neuroleptic-nonresponsive schizophrenic patients. *Schizophrenia Research, 77*(1), 35-41. doi:https://doi.org/10.1016/j.schres.2005.01.008

Joober, R., Rouleau, G. A., Lal, S., Dixon, M., O'Driscoll, G., Palmour, R., . . . Benkelfat, C. (2002). Neuropsychological impairments in neuroleptic-responder vs. -nonresponder schizophrenic patients and healthy volunteers. *Schizophr Res, 53*(3), 229-238.

Justice, A. E., Winkler, T. W., Feitosa, M. F., Graff, M., Fisher, V. A., Young, K., . . . Cupples, L. A. (2017). Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nature Communications, 8*, 14977. doi:10.1038/ncomms14977

Kahn, R. S., Sommer, I. E., Murray, R. M., Meyer-Lindenberg, A., Weinberger, D. R., Cannon, T. D., . . . Insel, T. R. (2015). Schizophrenia. *Nature Reviews Disease Primers, 1*(1), 15067. doi:10.1038/nrdp.2015.67

Kampman, O., Illi, A., Hänninen, K., Katila, H., Anttila, S., Rontu, R., . . . Lehtimäki, T. (2006). RGS4 genotype is not associated with antipsychotic medication response in schizophrenia. *Journal of Neural Transmission, 113*(10), 1563-1568. doi:10.1007/s00702-006-0445-2

Kapur, S., Agid, O., Mizrahi, R., & Li, M. (2006). How antipsychotics work-from receptors to reality. *NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics, 3*(1), 10-21. doi:10.1016/j.nurx.2005.12.003

Kapur, S., Phillips, A. G., & Insel, T. R. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry, 17*(12), 1174-1179. doi:10.1038/mp.2012.105

Kay, S. R., Fiszbein, A., & Opler, L. A. J. S. b. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *13*(2), 261-276.

Kennedy, J. L., Altar, C. A., Taylor, D. L., Degtiar, I., & Hornberger, J. C. (2014). The social and economic burden of treatment-resistant schizophrenia: a systematic literature review. *International clinical psychopharmacology, 29*(2), 63-76.

Kenny, D. A. (2015). Measuring Model Fit. Retrieved from http://davidakenny.net/cm/fit.htm

Kim, J. S., Park, C. M., Choi, J. A., Park, E., Tchoe, H. J., Choi, M., . . . Lee, S. H. (2017). The association between season of birth, age at onset, and clozapine use in schizophrenia. *Acta Psychiatr Scand, 136*(5), 445-454. doi:10.1111/acps.12776

Kinon, B. J., Chen, L., Ascher-Svanum, H., Stauffer, V. L., Kollack-Walker, S., Zhou, W., . . . Kane, J. M. (2010). Early Response to Antipsychotic Drug Therapy as a Clinical Marker of Subsequent Response in the Treatment of Schizophrenia. *Neuropsychopharmacology, 35*(2), 581-590. doi:10.1038/npp.2009.164

Kirkbride, J. B., Fearon, P., Morgan, C., Dazzan, P., Morgan, K., Tarrant, J., . . . Jones, P. B. (2006). Heterogeneity in incidence rates of schizophrenia and other psychotic syndromes: findings from the 3-center AeSOP study. *Arch Gen Psychiatry, 63*(3), 250-258. doi:10.1001/archpsyc.63.3.250

Koga, A., Bani-Fatemi, A., Hettige, N., Borlido, C., Zai, C., Strauss, J., . . . De Luca, V. (2017). GWAS analysis of treatment resistant schizophrenia: interaction effect of childhood trauma. *Pharmacogenomics, 18*(7), 663-671.

Kolenic, M., Franke, K., Hlinka, J., Matejka, M., Capkova, J., Pausova, Z., . . . Hajek, T. (2018). Obesity, dyslipidemia and brain age in first-episode psychosis. *J Psychiatr Res, 99*, 151-158. doi:10.1016/j.jpsychires.2018.02.012

Kontopantelis, E., Sperrin, M., Mamas, M. A., & Buchan, I. E. (2018). Investigating heterogeneity of effects and associations using interaction terms. *Journal of Clinical Epidemiology, 93*, 79-83. doi:https://doi.org/10.1016/j.jclinepi.2017.09.012

Krapohl, E., Euesden, J., Zabaneh, D., Pingault, J. B., Rimfeld, K., von Stumm, S., . . . Plomin, R. (2016). Phenome-wide analysis of genome-wide polygenic scores. *Mol Psychiatry, 21*(9), 1188-1193. doi:10.1038/mp.2015.126

Krapohl, E., Patel, H., Newhouse, S., Curtis, C. J., von Stumm, S., Dale, P. S., . . . Plomin, R. (2018). Multi-polygenic score approach to trait prediction. *Mol Psychiatry, 23*(5), 1368.

Krause, M., Zhu, Y., Huhn, M., Schneider-Thoma, J., Bighelli, I., Nikolakopoulou, A., & Leucht, S. (2018). Antipsychotic drugs for patients with schizophrenia and predominant or prominent negative symptoms: a systematic review and meta-analysis. *Eur Arch Psychiatry Clin Neurosci, 268*(7), 625-639. doi:10.1007/s00406-018-0869-3

Kravariti, E., Demjaha, A., Zanelli, J., Ibrahim, F., Wise, C., MacCabe, J. H., . . . Murray, R. M. (2018). Neuropsychological function at first episode in treatment-resistant psychosis: findings from the AESOP-10 study. *Psychological Medicine*, 1-11. doi:10.1017/s0033291718002957

Krebs, M. O., Guillin, O., Bourdel, M. C., Schwartz, J. C., Olie, J. P., Poirier, M. F., & Sokoloff, P. (2000). Brain Derived Neurotrophic Factor (BDNF) gene variants association with age at onset and therapeutic response in schizophrenia. *Mol Psychiatry, 5*(5), 558-562. doi:10.1038/sj.mp.4000749

Krebs, M. O., Sautel, F., Bourdel, M. C., Sokoloff, P., Schwartz, J. C., Olié, J. P., . . . Poirier, M. F. (1998). Dopamine D3 receptor gene variants and substance abuse in schizophrenia. *Mol Psychiatry, 3*(4), 337-341. doi:10.1038/sj.mp.4000411

Krivoy, A., Hochman, E., Sendt, K.-V., Hollander, S., Vilner, Y., Selakovic, M., . . . Taler, M. (2018). Association between serum levels of glutamate and neurotrophic factors and response to clozapine treatment. *Schizophrenia Research, 192*, 226-231. doi:https://doi.org/10.1016/j.schres.2017.05.040

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26): Springer.

Kumra, S., Kranzler, H., Gerbino-Rosen, G., Kester, H. M., DeThomas, C., Kafantaris, V., . . . Kane, J. M. (2008). Clozapine and "High-Dose" Olanzapine in Refractory Early-Onset Schizophrenia: A 12-Week Randomized and Double-Blind Comparison. *Biol Psychiatry, 63*(5), 524-529. doi:http://dx.doi.org/10.1016/j.biopsych.2007.04.043

Lahdelma, L., Ahokas, A., Andersson, L. C., Huttunen, M., Sarna, S., & Koskimies, S. (1998). Association between HLA-A1 allele and schizophrenia gene(s) in patients refractory to conventional neuroleptics but responsive to clozapine medication. *Tissue Antigens, 51*(2), 200-203. doi:10.1111/j.1399-0039.1998.tb02965.x

Lally, J., Ajnakina, O., Di Forti, M., Trotta, A., Demjaha, A., Kolliakou, A., . . . Murray, R. (2016a). Two distinct patterns of treatment resistance: Clinical predictors of treatment resistance in first-episode schizophrenia spectrum psychoses. *Psychological Medicine, 46*(15), 3231-3240. doi:http://dx.doi.org/10.1017/S0033291716002014

Lally, J., Gaughran, F., Timms, P., & Curran, S. R. (2016b). Treatment-resistant schizophrenia: Current insights on the pharmacogenomics of antipsychotics. *Pharmacogenomics and Personalized Medicine, 9*, 117-129. Retrieved from https://openathens.ovid.com/secure-ssl/home.oa?idpselect=https://kclidp.kcl.ac.uk/idp/shibboleth&entityID=https://kclidp.kcl.ac.uk/idp/shibboleth&?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=emexa&AN=613371854

http://sfx.kcl.ac.uk/kings?sid=OVID:embase&id=pmid:&id=doi:10.2147%2FPGPM.S115741&genre=article&atitle=Treatment-resistant+schizophrenia%3A+Current+insights+on+the+pharmacogenomi

cs+of+antipsychotics&title=Pharmacogenomics+and+Personalized+Medicine&issn=1178-7066&date=2016&volume=9&issue=&spage=117&aulast=Lally+J.&isbn=&__char_set=utf8

Lally, J., & MacCabe, J. H. (2015). Antipsychotic medication in schizophrenia: a review. *British medical bulletin, 114*(1), 169-179.

Lambert, C. G., & Black, L. J. (2012). Learning from our GWAS mistakes: from experimental design to scientific method. *Biostatistics, 13*(2), 195-203. doi:10.1093/biostatistics/kxr055

Lange, E. H., Nesvåg, R., Ringen, P. A., Hartberg, C. B., Haukvik, U. K., Andreassen, O. A., . . . Agartz, I. (2014). One year follow-up of alcohol and illicit substance use in first-episode psychosis: Does gender matter? *Comprehensive Psychiatry, 55*(2), 274-282. doi:https://doi.org/10.1016/j.comppsych.2013.08.018

Lawrence, R. E., & Appelbaum, P. S. (2011). Genetic Testing in Psychiatry: A Review of Attitudes and Beliefs. *Psychiatry: Interpersonal and Biological Processes, 74*(4), 315-331. doi:10.1521/psyc.2011.74.4.315

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., . . . Social Science Genetic Association, C. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet, 50*(8), 1112-1121. doi:10.1038/s41588-018-0147-3

Lee, S. H., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2012). A better coefficient of determination for genetic profile analysis. *Genet Epidemiol, 36*(3), 214-224.

Lee, Y. H., Bang, H., & Kim, D. J. (2016). How to Establish Clinical Prediction Models. *Endocrinology and metabolism (Seoul, Korea), 31*(1), 38-44. doi:10.3803/EnM.2016.31.1.38

Legge, S. E., Dennison, C. A., Pardiñas, A. F., Rees, E., Lynham, A. J., Hopkins, L., . . . Walters, J. T. R. (2019). Clinical indicators of treatment-resistant psychosis. *Br J Psychiatry*, 1-8. doi:10.1192/bjp.2019.120

Lencz, T., Guha, S., Liu, C., Rosenfeld, J., Mukherjee, S., DeRosse, P., . . . Badner, J. A. (2013). Genome-wide association study implicates NDST3 in schizophrenia and bipolar disorder. *Nature Communications, 4*.

Lester, H., & Wilson, S. (1999). Practical problems in recruiting patients with schizophrenia into randomised controlled trials. *BMJ, 318*(7190), 1075.

Leucht, S., Cipriani, A., Spineli, L., Mavridis, D., Orey, D., Richter, F., . . . Davis, J. M. (2013). Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet, 382*(9896), 951-962. doi:10.1016/s0140-6736(13)60733-3

Leucht, S., Corves, C., Arbter, D., Engel, R. R., Li, C., & Davis, J. M. (2009). Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *The Lancet, 373*(9657), 31-41. doi:https://doi.org/10.1016/S0140-6736(08)61764-X

Leucht, S., Tardy, M., Komossa, K., Heres, S., Kissling, W., Salanti, G., & Davis, J. M. (2012). Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. *Lancet, 379*(9831), 2063-2071. doi:10.1016/s0140-6736(12)60239-6

Levine, S. Z., & Rabinowitz, J. (2010). Trajectories and antecedents of treatment response over time in early-episode psychosis. *Schizophr Bull, 36*(3), 624-632. doi:10.1093/schbul/sbn120

Lewis, C. M., & Vassos, E. (2017). Prospects for using risk scores in polygenic medicine. *Genome medicine, 9*(1), 96.

Li, J., & Meltzer, H. Y. (2014). A genetic locus in 7p12.2 associated with treatment resistant schizophrenia. *Schizophrenia Research, 159*(2), 333-339. doi:https://doi.org/10.1016/j.schres.2014.08.018

Li, X., Sundquist, J., & Sundquist, K. (2007). Age-specific familial risks of psychotic disorders and schizophrenia: a nation-wide epidemiological study from Sweden. *Schizophr Res, 97*(1-3), 43-50. doi:10.1016/j.schres.2007.09.027

Li, Z., Chen, J., Yu, H., He, L., Xu, Y., Zhang, D., . . . Shi, Y. (2017). Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat Genet, 49*(11), 1576-1583. doi:10.1038/ng.3973

Lieberman, J. A., Stroup, T. S., McEvoy, J. P., Swartz, M. S., Rosenheck, R. A., Perkins, D. O., . . . Hsiao, J. K. (2005). Effectiveness of Antipsychotic Drugs in Patients with Chronic Schizophrenia. *New England Journal of Medicine, 353*(12), 1209-1223. doi:10.1056/NEJMoa051688

Lin, A., Kenis, G., Bignotti, S., Tura, G. J. B., De Jong, R., Bosmans, E., . . . Maes, M. (1998). The inflammatory response system in treatment-resistant schizophrenia: increased serum interleukin-6. *Schizophrenia Research, 32*(1), 9-15. doi:https://doi.org/10.1016/S0920-9964(98)00034-6

Lu, F., & Petkova, E. (2014). A comparative study of variable selection methods in the context of developing psychiatric screening instruments. *Stat Med, 33*(3), 401-421. doi:10.1002/sim.5937

Lyngstad, S. H., Gardsjord, E. S., Simonsen, C., Engen, M. J., Romm, K. L., Melle, I., & Færden, A. (2018). Consequences of persistent depression and apathy in first-episode psychosis — A one-year follow-up study. *Comprehensive Psychiatry, 86*, 60-66. doi:https://doi.org/10.1016/j.comppsych.2018.07.015

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods, 1*(2), 130.

Malla, A., Norman, R., Schmitz, N., Manchanda, R., BÉChard-Evans, L., Takhar, J., & Haricharan, R. A. J. (2006). Predictors of rate and time to remission in first-episode psychosis: a two-year outcome study. *Psychological Medicine, 36*(5), 649-658. doi:10.1017/S0033291706007379

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics (Oxford, England), 26*(22), 2867-2873. doi:10.1093/bioinformatics/btq559

Marees, A. T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., & Derks, E. M. (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International Journal of Methods in Psychiatric Research, 27*(2), e1608. doi:10.1002/mpr.1608

Martin, A. K., & Mowry, B. (2016). Increased rare duplication burden genomewide in patients with treatment-resistant schizophrenia. *Psychological Medicine, 46*(3), 469-476. doi:10.1017/S0033291715001701

Martinho Jr, E., Michelon, L., Ayres, A. M., Scazufca, M., Menezes, P. R., Schaufelberger, M. S., . . . Busatto Filho, G. (2012). BDNF gene polymorphism, cognition and symptom severity in a Brazilian population-based sample of first-episode psychosis subjects. *Revista Brasileira de Psiquiatria, 34*, 219-232.

Maxwell, M. E. (1992). Family Interview for Genetic Studies (FIGS): a manual for FIGS. *Bethesda, MD: Clinical Neurogenetics Branch, Intramural Research Program, National Institute of Mental Health*.

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., . . . Marchini, J. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet, 48*, 1279. doi:10.1038/ng.3643
https://www.nature.com/articles/ng.3643#supplementary-information

McCutcheon, R., Beck, K., Bloomfield, M. A., Marques, T. R., Rogdaki, M., & Howes, O. D. (2015). Treatment resistant or resistant to treatment? Antipsychotic plasma levels in patients with poorly controlled psychotic symptoms. *Journal of Psychopharmacology, 29*(8), 892-897. doi:10.1177/0269881115576688

McGrath, J. J., Saha, S., Al-Hamzawi, A. O., Alonso, J., Andrade, L., Borges, G., . . . Kessler, R. C. (2016). Age of Onset and Lifetime Projected Risk of Psychotic Experiences: Cross-National Data From the World Mental Health Survey. *Schizophr Bull, 42*(4), 933-941. doi:10.1093/schbul/sbw011

McGuffin, P., Farmer, A., & Harvey, I. (1991). A Polydiagnostic Application of Operational Criteria in Studies of Psychotic Illness: Development and Reliability of the OPCRIT System. *JAMA Psychiatry, 48*(8), 764-770. doi:10.1001/archpsyc.1991.01810320088015

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., . . . Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology, 17*(1), 122. doi:10.1186/s13059-016-0974-4

Meged, S., Stein, D., Sitrota, P., Melamed, Y., Elizur, A., Shmuelian, I., & Gazit, E. (1999). Human leukocyte antigen typing, response to neuroleptics, and clozapine-induced agranulocytosis in jewish Israeli schizophrenic patients. *International clinical psychopharmacology, 14*(5), 305-312. Retrieved from http://europepmc.org/abstract/MED/10529073

Melicher, T., Horacek, J., Hlinka, J., Spaniel, F., Tintera, J., Ibrahim, I., . . . Hoschl, C. (2015). White matter changes in first episode psychosis and their relation to the size of sample studied: a DTI study. *Schizophr Res, 162*(1-3), 22-28. doi:10.1016/j.schres.2015.01.029

Meltzer, H. Y., Rabinowitz, J., Lee, M. A., Cola, P. A., Ranjan, R., Findling, R. L., & Thompson, P. A. (1997). Age at onset and gender of schizophrenic patients in relation to neuroleptic resistance. *American Journal of Psychiatry, 154*(4), 475-482. doi:10.1176/ajp.154.4.475

Menezes, P. R., Scazufca, M., Busatto, G. F., Coutinho, L. M., McGuire, P. K., & Murray, R. M. (2007). Incidence of first-contact psychosis in Sao Paulo, Brazil. *The British Journal of Psychiatry, 191*(S51), s102-s106.

Mikolas, P., Melicher, T., Skoch, A., Matejka, M., Slovakova, A., Bakstein, E., . . . Spaniel, F. (2016). Connectivity of the anterior insula differentiates participants with first-episode schizophrenia spectrum disorders from controls: a machine-learning study. *Psychological Medicine, 46*(13), 2695-2704. doi:10.1017/S0033291716000878

Minatogawa-Chang, T. M., Schaufelberger, M. S., Ayres, A. M., Duran, F. L., Gutt, E. K., Murray, R. M., . . . Scazufca, M. (2009). Cognitive performance is related to cortical grey matter volumes in early stages of schizophrenia: a population-based study of first-episode psychosis. *Schizophrenia Research, 113*(2-3), 200-209.

Moreno-Küstner, B., Martín, C., & Pastor, L. (2018). Prevalence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. *PLoS One, 13*(4), e0195687. doi:10.1371/journal.pone.0195687

Morgan, C., Dazzan, P., Morgan, K., Jones, P., Harrison, G., Leff, J., . . . Fearon, P. (2006). First episode psychosis and ethnicity: initial findings from the AESOP study. *World Psychiatry, 5*(1), 40-46.

Morgan, C., Lappin, J., Heslin, M., Donoghue, K., Lomas, B., Reininghaus, U., . . . Dazzan, P. (2014). Reappraising the long-term course and outcome of psychotic disorders: the AESOP-10 study. *Psychol Med, 44*(13), 2713-2726. doi:10.1017/s0033291714000282

Morrison, A. P., Pyle, M., Gumley, A., Schwannauer, M., Turkington, D., MacLennan, G., . . . French, P. (2018). Cognitive behavioural therapy in clozapine-resistant schizophrenia (FOCUS): an assessor-blinded, randomised controlled trial. *The Lancet Psychiatry, 5*(8), 633-643.

Mortimer, A., Singh, P., Shepherd, C., & Puthiryackal, J. (2010). Clozapine for Treatment-Resistant Schizophrenia: National Institute of Clinical Excellence (NICE) Guidance in the Real World. *Clinical Schizophrenia & Related Psychoses, 4*(1), 49-55. doi:10.3371/csrp.4.1.4

Mouchlianitis, E., Bloomfield, M. A., Law, V., Beck, K., Selvaraj, S., Rasquinha, N., . . . Stone, J. (2015). Treatment-resistant schizophrenia patients show elevated anterior cingulate cortex glutamate compared to treatment-responsive. *Schizophr Bull, 42*(3), 744-752.

Murru, A., & Carpiniello, B. (2018). Duration of untreated illness as a key to early intervention in schizophrenia: A review. *Neuroscience Letters, 669*, 59-67. doi:https://doi.org/10.1016/j.neulet.2016.10.003

Narvaez, J. M., Twamley, E. W., McKibbin, C. L., Heaton, R. K., & Patterson, T. L. (2008). Subjective and objective quality of life in schizophrenia. *Schizophrenia Research, 98*(1-3), 201-208.

National Institute for Health and Care Excellence. (2014). *Psychosis and Schizophrenia in Adults: Treatment and Management (Clinical Guideline 178)*. London: Royal College of Psychiatrists.

Nelson, M. T., Seal, M. L., Pantelis, C., & Phillips, L. J. (2013). Evidence of a dimensional relationship between schizotypy and schizophrenia: A systematic review. *Neuroscience & Biobehavioral Reviews, 37*(3), 317-327. doi:https://doi.org/10.1016/j.neubiorev.2013.01.004

Nguyen, C. D., Carlin, J. B., & Lee, K. J. (2017). Model checking in multiple imputation: an overview and case study. *Emerg Themes Epidemiol, 14*, 8. doi:10.1186/s12982-017-0062-6

NICE. (2014). Psychosis and Schizophrenia in Adults: Treatment and Management: Updated Edition 2014. In *National Institute for Health and Clinical Excellence: Guidance* London: National Collaborating Centre for Mental Health.

Nielsen, J., Dahm, M., Lublin, H., & Taylor, D. (2010). Psychiatrists' attitude towards and knowledge of clozapine treatment. *Journal of Psychopharmacology, 24*(7), 965-971.

Niemi, M. E., Martin, H. C., Rice, D. L., Gallone, G., Gordon, S., Kelemen, M., . . . Barrett, J. C. (2018). Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *bioRxiv*. doi:10.1101/309070

Nordstroem, A.-L., Talbot, D., Bernasconi, C., Berardo, C. G., & Lalonde, J. (2017). Burden of illness of people with persistent symptoms of schizophrenia: a multinational cross-sectional study. *International Journal of Social Psychiatry, 63*(2), 139-150.

O'Donoghue, B., Lyne, J., Madigan, K., Lane, A., Turner, N., O'Callaghan, E., & Clarke, M. (2015). Environmental factors and the age at onset in first episode psychosis. *Schizophrenia Research, 168*(1), 106-112. doi:https://doi.org/10.1016/j.schres.2015.07.004

Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., . . . Benjamin, D. J. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature, 533*(7604), 539-542. doi:10.1038/nature17671

Ota, V. K., Spíndola, L. N., Gadelha, A., Santos Filho, A. F. d., Santoro, M. L., Christofolini, D. M., . . . Belangero, S. I. (2012). DRD1 rs4532 polymorphism: A potential pharmacogenomic marker for treatment response to antipsychotic drugs. *Schizophrenia Research, 142*(1), 206-208. doi:https://doi.org/10.1016/j.schres.2012.08.003

Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological Reports, 10*(3), 799-812.

Palk, A. C., Dalvie, S., de Vries, J., Martin, A. R., & Stein, D. J. (2019). Potential use of clinical polygenic risk scores in psychiatry – ethical implications and communicating high polygenic risk. *Philosophy, Ethics, and Humanities in Medicine, 14*(1), 4. doi:10.1186/s13010-019-0073-8

Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., . . . Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet, 50*(3), 381-389. doi:10.1038/s41588-018-0059-2

Pavlou, M., Ambler, G., Seaman, S., De Iorio, M., & Omar, R. Z. (2016). Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. *Statistics in Medicine, 35*(7), 1159-1177. doi:10.1002/sim.6782

Pelayo-Teran, J. M., Perez-Iglesias, R., Ramirez-Bonilla, M., Gonzalez-Blanch, C., Martinez-Garcia, O., Pardo-Garcia, G., . . . Crespo-Facorro, B. (2008). Epidemiological factors associated with treated incidence of first-episode non-affective psychosis in Cantabria: insights from the Clinical Programme on Early Phases of Psychosis. *Early Interv Psychiatry, 2*(3), 178-187. doi:10.1111/j.1751-7893.2008.00074.x

Peloso, G. M., & Lunetta, K. L. (2011). Choice of population structure informative principal components for adjustment in a case-control study. *BMC Genetics, 12*, 64-64. doi:10.1186/1471-2156-12-64

Pettersson, F. H., Anderson, C. A., Clarke, G. M., Barrett, J. C., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2009). Marker selection for genetic

case-control association studies. *Nature protocols, 4*(5), 743-752. doi:10.1038/nprot.2009.38

Pickrell, J. K., Berisa, T., Liu, J. Z., Segurel, L., Tung, J. Y., & Hinds, D. A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet, 48*(7), 709-717. doi:10.1038/ng.3570

http://www.nature.com/ng/journal/v48/n7/abs/ng.3570.html#supplementary-information

Pinheiro, D. S., Santos, R. d. S., de Brito, R. B., Cruz, A. H. d. S., Ghedini, P. C., & Reis, A. A. (2017). GSTM1/GSTT1 double-null genotype increases risk of treatment-resistant schizophrenia: A genetic association study in Brazilian patients. *PLoS One, 12*(8), e0183812.

Pisanu, C., & Squassina, A. (2019). Treatment-Resistant Schizophrenia: Insights From Genetic Studies and Machine Learning Approaches. *Front Pharmacol, 10*, 617. doi:10.3389/fphar.2019.00617

Power, R. A., Tansey, K. E., Buttenschon, H. N., Cohen-Woods, S., Bigdeli, T., Hall, L. S., . . . Lewis, C. M. (2017). Genome-wide Association for Major Depression Through Age at Onset Stratification: Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. *Biol Psychiatry, 81*(4), 325-335. doi:10.1016/j.biopsych.2016.05.010

Price, A. L., Weale, M. E., Patterson, N., Myers, S. R., Need, A. C., Shianna, K. V., . . . Taylor, K. D. (2008). Long-range LD can confound genome scans in admixed populations. *American journal of human genetics, 83*(1), 132.

Pruessner, M., Cullen, A. E., Aas, M., & Walker, E. F. (2017). The neural diathesis-stress model of schizophrenia revisited: An update on recent findings considering illness stage and neurobiological and methodological complexities. *Neuroscience & Biobehavioral Reviews, 73*, 191-218. doi:https://doi.org/10.1016/j.neubiorev.2016.12.013

Quteineh, L., Vandenberghe, F., Saigi Morgui, N., Delacretaz, A., Choong, E., Gholam-Rezaee, M., . . . Eap, C. B. (2015). Impact of HSD11B1 polymorphisms on BMI and components of the metabolic syndrome in patients receiving psychotropic treatments. *Pharmacogenet Genomics, 25*(5), 246-258. doi:10.1097/fpc.0000000000000131

R Core Team. (2009). R: A language and environment for statistical computing. Retrieved from https://www.R-project.org/

Rajji, T. K., Ismail, Z., & Mulsant, B. H. (2009). Age at onset and cognition in schizophrenia: meta-analysis. *Br J Psychiatry, 195*(4), 286-293. doi:10.1192/bjp.bp.108.060723

Ramasamy, A., Trabzuni, D., Guelfi, S., Varghese, V., Smith, C., Walker, R., . . . Weale, M. E. (2014). Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci, 17*(10), 1418-1428. doi:10.1038/nn.3801

http://www.nature.com/neuro/journal/v17/n10/abs/nn.3801.html#supplementary-information

Reed, E., Nunez, S., Kulp, D., Qian, J., Reilly, M. P., & Foulkes, A. S. (2015). A guide to genome-wide association analysis and post-analytic interrogation. *Statistics in Medicine, 34*(28), 3769-3792. doi:10.1002/sim.6605

Rees, E., Kirov, G., Sanders, A., Walters, J. T. R., Chambert, K. D., Shi, J., . . . Owen, M. J. (2014). Evidence that duplications of 22q11.2 protect against schizophrenia. *Mol Psychiatry, 19*(1), 37-40. doi:10.1038/mp.2013.156

Rees, E., O'Donovan, M. C., & Owen, M. J. (2015). Genetics of schizophrenia. *Current Opinion in Behavioral Sciences, 2*, 8-14. doi:https://doi.org/10.1016/j.cobeha.2014.07.001

Revier, C. J., Reininghaus, U., Dutta, R., Fearon, P., Murray, R. M., Doody, G. A., . . . Jones, P. B. (2015). Ten-Year Outcomes of First-Episode Psychoses in the MRC ÆSOP-10 Study. *The Journal of Nervous and Mental Disease, 203*(5), 379-386. doi:10.1097/NMD.0000000000000295

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychol Methods, 17*(3), 354.

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., . . . Koellinger, P. D. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science, 340*(6139), 1467-1471. doi:10.1126/science.1235488

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K.-H., Holmans, P. A., . . . Huang, H. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature, 511*(7510), 421.

Ripke, S., Sanders, A. R., Kendler, K. S., Levinson, D. F., Sklar, P., Holmans, P. A., . . . Gejman, P. V. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nat Genet, 43*, 969. doi:10.1038/ng.940

https://www.nature.com/articles/ng.940#supplementary-information

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* New York, NY: John Wiley & Sons.

Ruderfer, D. M., Charney, A. W., Readhead, B., Kidd, B. A., Kähler, A. K., Kenny, P. J., . . . Sklar, P. (2016). Polygenic overlap between schizophrenia risk and antipsychotic response: a genomic medicine approach. *The Lancet Psychiatry, 3*(4), 350-357. doi:https://doi.org/10.1016/S2215-0366(15)00553-2

Ruderfer, D. M., Ripke, S., McQuillin, A., Boocock, J., Stahl, E. A., Pavlides, J. M. W., . . . Kendler, K. S. (2018). Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell, 173*(7), 1705-1715.e1716. doi:10.1016/j.cell.2018.05.046

Sa, A. R., Hounie, A. G., Sampaio, A. S., Arrais, J., Miguel, E. C., & Elkis, H. (2009). Obsessive-compulsive symptoms and disorder in patients with schizophrenia treated with clozapine or haloperidol. *Comprehensive Psychiatry, 50*(5), 437-442.

Salvador, R., Radua, J., Canales-Rodriguez, E. J., Solanes, A., Sarro, S., Goikolea, J. M., . . . Pomarol-Clotet, E. (2017). Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS One, 12*(4), e0175683. doi:10.1371/journal.pone.0175683

Samanaite, R., Gillespie, A., Sendt, K.-V., McQueen, G., MacCabe, J. H., & Egerton, A. (2018). Biological Predictors of Clozapine Response: A Systematic Review. *Frontiers in Psychiatry, 9*, 327-327. doi:10.3389/fpsyt.2018.00327

Samara, M. T., Leucht, C., Leeflang, M. M., Anghelescu, I.-G., Chung, Y.-C., Crespo-Facorro, B., . . . Kane, J. M. (2015). Early improvement as a predictor of later response to antipsychotics in schizophrenia: a diagnostic test review. *American Journal of Psychiatry, 172*(7), 617-629.

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., . . . Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet, 50*(7), 912-919. doi:10.1038/s41588-018-0152-6

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychol Methods, 7*(2), 147-177.

Schaufelberger, M., Lappin, J., Duran, F., Rosa, P., Uchida, R., Santos, L., . . . Menezes, P. (2011). Lack of progression of brain abnormalities in first-episode psychosis: a longitudinal magnetic resonance imaging study. *Psychological Medicine, 41*(8), 1677-1689.

Setien-Suero, E., Martinez-Garcia, O., de la Foz, V. O., Vazquez-Bourgon, J., Correa-Ghisays, P., Ferro, A., . . . Ayesa-Arriola, R. (2018). Age of onset of Cannabis use and cognitive function in first-episode non-affective psychosis patients: Outcome at three-year follow-up. *Schizophr Res*. doi:10.1016/j.schres.2018.05.036

Shmueli, G. (2010). To explain or to predict? *Statistical science, 25*(3), 289-310.

Simeone, J. C., Ward, A. J., Rotella, P., Collins, J., & Windisch, R. (2015). An evaluation of variation in published estimates of schizophrenia prevalence from 1990─2013: a systematic literature review. *BMC Psychiatry, 15*(1), 193. doi:10.1186/s12888-015-0578-7

Siskind, D., McCartney, L., Goldschlager, R., & Kisely, S. (2016). Clozapine v. first- and second-generation antipsychotics in treatment-refractory schizophrenia: systematic review and meta-analysis. *Br J Psychiatry, 209*(5), 385-392. doi:10.1192/bjp.bp.115.177261

Siskind, D., Siskind, V., & Kisely, S. (2017). Clozapine response rates among people with treatment-resistant schizophrenia: data from a systematic review and meta-analysis. In: SAGE Publications Sage CA: Los Angeles, CA.

Smart, S. E., Kepinska, A. P., Murray, R. M., & MacCabe, J. H. (2019). Predictors of treatment resistant schizophrenia: a systematic review of prospective observational studies. *Psychol Med*, 1-10. doi:10.1017/s0033291719002083

Sniekers, S., Stringer, S., Watanabe, K., Jansen, P. R., Coleman, J. R. I., Krapohl, E., . . . Posthuma, D. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat Genet, 49*(7), 1107-1112. doi:10.1038/ng.3869

http://www.nature.com/ng/journal/v49/n7/abs/ng.3869.html#supplementary-information

Sorensen, H. J., Foldager, L., Roge, R., Pristed, S. G., Andreasen, J. T., & Nielsen, J. (2014). An association between autumn birth and clozapine treatment in patients with schizophrenia: a population-based analysis. *Nordic Journal of Psychiatry, 68*(6), 428-432. doi:10.3109/08039488.2013.854408

Spaniel, F., Tintera, J., Rydlo, J., Ibrahim, I., Kasparek, T., Horacek, J., . . . Hajek, T. (2016). Altered Neural Correlate of the Self-Agency Experience in First-Episode Schizophrenia-Spectrum Patients: An fMRI Study. *Schizophr Bull, 42*(4), 916-925. doi:10.1093/schbul/sbv188

Stanaway, I. B., Hall, T. O., Rosenthal, E. A., Palmer, M., Naranbhai, V., Knevel, R., . . . Crosslin, D. R. (2019). The eMERGE genotype set of 83,717 subjects imputed to ~40 million variants genome wide and association with

the herpes zoster medical record phenotype. *Genet Epidemiol, 43*(1), 63-81. doi:10.1002/gepi.22167

Stekhoven, D. J., & Buhlmann, P. (2012). MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics (Oxford, England), 28*(1), 112-118. doi:10.1093/bioinformatics/btr597

Steyerberg, E. W. (2008). *Clinical prediction models: a practical approach to development, validation, and updating*: Springer Science & Business Media.

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., . . . Kattan, M. W. (2010). Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology, 21*(1), 128-138. doi:10.1097/EDE.0b013e3181c30fb2

Stone-Romero, E. F., & Anderson, L. E. (1994). Relative power of moderated multiple regression and the comparison of subgroup correlation coefficients for detecting moderating effects. *Journal of Applied Psychology, 79*(3), 354-359. doi:10.1037/0021-9010.79.3.354

Streiner, D. L. (2003). Diagnosing tests: Using and misusing diagnostic and screening tests. *Journal of personality assessment, 81*(3), 209-219.

Streiner, D. L. (2005). Finding our way: an introduction to path analysis. *Can J Psychiatry, 50*(2), 115-122. doi:10.1177/070674370505000207

Sullivan, P. F., Agrawal, A., Bulik, C. M., Andreassen, O. A., Børglum, A. D., Breen, G., . . . Gelernter, J. (2017). Psychiatric genomics: an update and an agenda. *American Journal of Psychiatry, 175*(1), 15-27.

Suvisaari, J., Mantere, O., Keinanen, J., Mantyla, T., Rikandi, E., Lindgren, M., . . . Raij, T. T. (2018). Is It Possible to Predict the Future in First-Episode Psychosis? *Front Psychiatry, 9*, 580. doi:10.3389/fpsyt.2018.00580

Szymanski, S., Lieberman, J. A., Alvir, J. M., Mayerhoff, D., Loebel, A., Geisler, S., . . . et al. (1995). Gender differences in onset of illness, treatment response, course, and biologic indexes in first-episode schizophrenic patients. *Am J Psychiatry, 152*(5), 698-703. doi:10.1176/ajp.152.5.698

Takeuchi, H., Thiyanavadivel, S., Fervaha, G., & Remington, G. (2017). Neurocognitive Benefits of Second-Generation Antipsychotics Versus Placebo: Insufficient Evidence Based on a Systematic Review. *Journal of clinical psychopharmacology, 37*(2), 274-276. doi:10.1097/jcp.0000000000000662

Tamminga, C. A., & Holcomb, H. H. (2005). Phenotype of schizophrenia: a review and formulation. *Mol Psychiatry, 10*(1), 27-39. doi:10.1038/sj.mp.4001563

Tarricone, I., Mimmi, S., Paparelli, A., Rossi, E., Mori, E., Panigada, S., . . . Berardi, D. (2012). First-episode psychosis at the West Bologna Community Mental Health Centre: results of an 8-year prospective study. *Psychol Med, 42*(11), 2255-2264. doi:10.1017/s0033291712000335

Taylor, D. M. (2017). Clozapine for Treatment-Resistant Schizophrenia: Still the Gold Standard? *CNS Drugs, 31*(3), 177-180. doi:10.1007/s40263-017-0411-6

Teo, C., Zai, C., Borlido, C., Tomasetti, C., Strauss, J., Shinkai, T., . . . De Luca, V. (2012). Analysis of treatment-resistant schizophrenia and 384 markers from candidate genes. *Pharmacogenet Genomics, 22*(11), 807-811.

Terzić, T., Kastelic, M., Dolžan, V., & Plesničar, B. K. (2015). Influence of 5-HT1A and 5-HTTLPR genetic variants on the schizophrenia symptoms and

occurrence of treatment-resistant schizophrenia. *Neuropsychiatric disease and treatment, 11*, 453-459. doi:10.2147/NDT.S76494

Tiihonen, J., Mittendorfer-Rutz, E., Majak, M., & et al. (2017a). Real-world effectiveness of antipsychotic treatments in a nationwide cohort of 29 823 patients with schizophrenia. *JAMA Psychiatry*. doi:10.1001/jamapsychiatry.2017.1322

Tiihonen, J., Mittendorfer-Rutz, E., Majak, M., Mehtälä, J., Hoti, F., Jedenius, E., . . . Taipale, H. (2017b). Real-World Effectiveness of Antipsychotic Treatments in a Nationwide Cohort of 29 823 Patients With Schizophrenia. *JAMA Psychiatry, 74*(7), 686-693. doi:10.1001/jamapsychiatry.2017.1322

Turkington, A., Mulholland, C. C., Rushe, T. M., Anderson, R., McCaul, R., Barrett, S. L., . . . Cooper, S. J. (2018). Impact of persistent substance misuse on 1-year outcome in first-episode psychosis. *British Journal of Psychiatry, 195*(3), 242-248. doi:10.1192/bjp.bp.108.057471

Turner, S., Armstrong, L. L., Bradford, Y., Carlson, C. S., Crawford, D. C., Crenshaw, A. T., . . . Ritchie, M. D. (2011). Quality control procedures for genome-wide association studies. *Current protocols in human genetics, Chapter 1*, Unit1.19-Unit11.19. doi:10.1002/0471142905.hg0119s68

Üçok, A., Çıkrıkçılı, U., Ergül, C., Tabak, Ö., Salaj, A., Karabulut, S., & Correll, C. U. (2016). Correlates of Clozapine Use after a First Episode of Schizophrenia: Results From a Long-term Prospective Study. *CNS Drugs, 30*(10), 997-1006. doi:10.1007/s40263-016-0358-z

Ucok, A., Polat, A., Cakir, S., & Genc, A. (2006). One year outcome in first episode schizophrenia. Predictors of relapse. *Eur Arch Psychiatry Clin Neurosci, 256*(1), 37-43. doi:10.1007/s00406-005-0598-2

Ucok, A., Polat, A., Genc, A., Cakir, S., & Turan, N. (2004). Duration of untreated psychosis may predict acute treatment response in first-episode schizophrenia. *J Psychiatr Res, 38*(2), 163-168.

Ucok, A., Serbest, S., & Kandemir, P. E. (2011). Remission after first-episode schizophrenia: results of a long-term follow-up. *Psychiatry Res, 189*(1), 33-37. doi:10.1016/j.psychres.2010.11.013

Üçok, A., Yazici, K., Mete, L., Kültür, S., Göğüs, A., Erkoç, Ş., & Alptekin, K. (2007). Factors associated with dropout and noncompliance in patients with schizophrenia: Results of a one-year follow-up. *Clinical Schizophrenia & Related Psychoses, 1*(2), 161-167.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *2011, 45*(3), 67. doi:10.18637/jss.v045.i03

van de Bilt, M. T., Prado, C. M., Ojopi, E. P. B., Sousa, R. T., Loch, A. A., Zanetti, M. V., . . . Gattaz, W. F. (2015). Cytochrome P450 genotypes are not associated with refractoriness to antipsychotic treatment. *Schizophrenia Research, 168*(1), 587-588. doi:https://doi.org/10.1016/j.schres.2015.08.002

Van Erp, T. G., Preda, A., Nguyen, D., Faziola, L., Turner, J., Bustillo, J., . . . Voyvodic, J. J. S. r. (2014). Converting positive and negative symptom scores between PANSS and SAPS/SANS. *152*(1), 289-294.

Vandenberghe, F., Gholam-Rezaee, M., Saigi-Morgui, N., Delacretaz, A., Choong, E., Solida-Tozzi, A., . . . Eap, C. B. (2015). Importance of early weight changes to predict long-term weight gain during psychotropic drug treatment. *J Clin Psychiatry, 76*(11), e1417-1423. doi:10.4088/JCP.14m09358

Vickers, A. J., Kramer, B. S., & Baker, S. G. (2006). Selecting patients for randomized trials: a systematic approach based on risk group. *Trials, 7*(1), 30.

Vilhjalmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindstrom, S., Ripke, S., . . . Price, A. L. (2015). Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *American journal of human genetics, 97*(4), 576-592. doi:10.1016/j.ajhg.2015.09.001

Vos, T., Barber, R. M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., . . . Murray, C. J. L. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet, 386*(9995), 743-800. doi:https://doi.org/10.1016/S0140-6736(15)60692-4

Warnez, S., & Alessi-Severini, S. (2014). Clozapine: a review of clinical practice guidelines and prescribing trends. *BMC Psychiatry, 14*(1), 102. doi:10.1186/1471-244X-14-102

Watson, P., Zhang, J.-P., Rizvi, A., Tamaiev, J., Birnbaum, M. L., & Kane, J. (2018). A meta-analysis of factors associated with quality of life in first episode psychosis. *Schizophrenia Research, 202*, 26-36. doi:https://doi.org/10.1016/j.schres.2018.07.013

Wheeler, A. J. (2008). Treatment Pathway and Patterns of Clozapine Prescribing for Schizophrenia in New Zealand. *Annals of Pharmacotherapy, 42*(6), 852-860. doi:10.1345/aph.1K662

Wils, R. S., Gotfredsen, D. R., Hjorthøj, C., Austin, S. F., Albert, N., Secher, R. G., . . . Nordentoft, M. (2017). Antipsychotic medication and remission of psychotic symptoms 10 years after a first-episode psychosis. *Schizophrenia Research, 182*, 42-48.

Wimberley, T., Gasse, C., Meier, S. M., Agerbo, E., MacCabe, J. H., & Horsdal, H. T. (2017). Polygenic Risk Score for Schizophrenia and Treatment-Resistant Schizophrenia. *Schizophr Bull, 43*(5), 1064-1069. doi:10.1093/schbul/sbx007

Wimberley, T., Pedersen, C. B., MacCabe, J. H., Støvring, H., Astrup, A., Sørensen, H. J., . . . Gasse, C. (2016a). Inverse association between urbanicity and treatment resistance in schizophrenia. *Schizophrenia Research, 174*(1), 150-155. doi:https://doi.org/10.1016/j.schres.2016.03.021

Wimberley, T., Stovring, H., Sorensen, H. J., Horsdal, H. T., MacCabe, J. H., & Gasse, C. (2016b). Predictors of treatment resistance in patients with schizophrenia: a population-based cohort study. *Lancet Psychiatry, 3*(4), 358-366. doi:10.1016/s2215-0366(15)00575-1

Wimberley, T., Støvring, H., Sørensen, H. J., Horsdal, H. T., MacCabe, J. H., & Gasse, C. (2016c). Predictors of treatment resistance in patients with schizophrenia: a population-based cohort study. *The Lancet Psychiatry, 3*(4), 358-366. doi:https://doi.org/10.1016/S2215-0366(15)00575-1

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*: Geneva: World Health Organization.

Wray, N. R., Lee, S. H., Mehta, D., Vinkhuyzen, A. A. E., Dudbridge, F., & Middeldorp, C. M. (2014). Research Review: Polygenic methods and their

application to psychiatric traits. *Journal of Child Psychology and Psychiatry, 55*(10), 1068-1087. doi:10.1111/jcpp.12295

Yoshimura, B., Yada, Y., So, R., Takaki, M., & Yamada, N. (2017). The critical treatment window of clozapine in treatment-resistant schizophrenia: Secondary analysis of an observational study. *Psychiatry Res, 250*, 65-70. doi:10.1016/j.psychres.2017.01.064

Yu, H., Yan, H., Li, J., Li, Z., Zhang, X., Ma, Y., . . . Yue, W. (2017). Common variants on 2p16.1, 6p22.1 and 10q24.32 are associated with schizophrenia in Han Chinese population. *Mol Psychiatry, 22*(7), 954-960. doi:10.1038/mp.2016.212

Yung, A., Phillips, L., McGorry, P., Ward, J., Donovan, K., & Thompson, K. (2002). Comprehensive assessment of at-risk mental states (CAARMS). *Melbourne, Australia, University of Melbourne, Department of Psychiatry, Personal Assessment and Crisis Evaluation Clinic*.

Zhang, J.-P., Lencz, T., Geisler, S., DeRosse, P., Bromet, E. J., & Malhotra, A. K. (2013). Genetic variation in BDNF is associated with antipsychotic treatment resistance in patients with schizophrenia. *Schizophrenia Research, 146*(1), 285-288. doi:https://doi.org/10.1016/j.schres.2013.01.020

Zhang, Z., Ma, P., Li, Q., Xiao, Q., Sun, H., Olasege, B. S., . . . Pan, Y. (2018). Exploring the Genetic Correlation Between Growth and Immunity Based on Summary Statistics of Genome-Wide Association Studies. *Frontiers in Genetics, 9*(393). doi:10.3389/fgene.2018.00393

Zheng, J., Erzurumluoglu, A. M., Elsworth, B. L., Kemp, J. P., Howe, L., Haycock, P. C., . . . Neale, B. M. (2017). LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics (Oxford, England), 33*(2), 272-279. doi:10.1093/bioinformatics/btw613

Zhu, Y., Li, C., Huhn, M., Rothe, P., Krause, M., Bighelli, I., . . . Leucht, S. (2017). How well do patients with a first episode of schizophrenia respond to antipsychotics: A systematic review and meta-analysis. *Eur Neuropsychopharmacol, 27*(9), 835-844. doi:10.1016/j.euroneuro.2017.06.011

Zimbron, J., Stahl, D., Hutchinson, G., Dazzan, P., Morgan, K., Doody, G. A., . . . MacCabe, J. H. (2014). Pre-morbid fertility in psychosis: Findings from the AESOP first episode study. *Schizophrenia Research, 156*(2), 168-173. doi:10.1016/j.schres.2014.04.007

Zondervan, K. T., & Cardon, L. R. (2007). Designing candidate gene and genome-wide case-control association studies. *Nature protocols, 2*(10), 2492-2501. doi:10.1038/nprot.2007.366

APPENDICES

14. Appendices

14.1. Appendix A

Table 15. Percentage of missingness stratified by outcome.

Variable	Outcome		
	Total	**Non-TR**	**TR**
Age of onset (years)	15.94	15.31	19.26
Age at baseline (years)	7.80	6.48	14.78
DUP (days)	28.38	26.46	38.52
Female (vs. male)	15.94	15.01	20.84
Body Mass Index	62.97	61.55	70.45
Being in a relationship (vs. single)	31.29	32.13	26.91
Ever being in a relationship (vs. never)	31.29	32.13	26.91
Living alone (vs. with others)	50.91	51.15	49.60
Living with non-parents (vs. with parents)	70.60	71.03	68.34
Living with non-family (vs. with family)	66.81	66.67	67.55
Living independent (vs. supported accommodation)	82.75	81.53	89.18
Employed (vs. unemployed)	75.03	75.20	74.14
Years in Education	69.17	68.47	72.82
Cannabis (vs. no cannabis)	45.42	43.07	57.78
Non-smoker (vs. smoker)	56.56	54.47	67.55
Non-drinker (vs. drinker)	54.49	51.41	70.71
PANSS Positive	71.49	70.23	78.10
PANSS Negative	71.70	70.43	78.36
PANSS General Psychopathology	71.87	70.63	78.36
PANSS Total Score	72.33	71.08	78.89
SAPS	60.14	57.78	72.56
SANS	60.35	58.03	72.56
GAF	74.91	75.45	72.03
BPRS Total Score	77.77	77.91	77.04
Family history of psychosis (vs. none)	83.80	83.48	85.49
Family history of mental health disorders (vs. none)	82.12	81.63	84.70
Highest education qualification	72.54	75.30	58.05
Ethnicity	32.22	33.68	24.54
Length of follow-up (years)	2.74	3.11	0.79

Table 16. Percentage missingness stratified by cohort.

Variable	Cohort											
	AESOP London	**Belfast**	**Bologna**	**GAP London**	**Istanbul**	**Lausanne**	**Oslo**	**Paris**	**Prague**	**Santander**	**Sao Paulo**	**UCL London**
Age of onset (years)	0.70	27.45	0	2.44	1.75	33.57	0	3.13	100	2.74	100	3.44
Age at baseline (years)	0	18.95	2.00	2.09	1.75	0.71	0	0	55.71	0.21	100	0.53
DUP (days)	8.39	34.64	100	100	1.75	33.21	0	6.25	56.43	2.74	100	1.59
Female (vs. male)	0	0	0	0	5.26	90.71	0	0	55.71	0.21	3.08	10.58
Body Mass Index	100	47.71	100	42.16	100	100	2.38	100	100	1.47	100	100
Being in a relationship (vs. single)	7.34	6.54	18.00	29.27	0	33.21	0	3.13	55.71	0.63	100	100
Ever being in a relationship (vs. never)	7.34	6.54	18.00	29.27	0	33.21	0	3.13	55.71	0.63	100	100
Living alone (vs. with others)	7.34	72.55	18.00	29.62	100	100	15.48	100	100	0.63	100	100
Living with non-parents (vs. with parents)	45.80	81.05	28.00	60.63	100	100	100	100	100	23.37	100	100
Living with non-family (vs. with family)	45.80	81.05	28.00	60.63	100	100	46.43	100	100	23.37	100	100
Living independent (vs. supported accommodation)	100	100	24.00	100	100	35.71	0.60	25.00	100	100	100	100
Employed (vs. unemployed)	100	13.73	18.00	31.01	3.51	100	19.64	3.13	100	100	100	100
Years in Education	100	37.91	18.00	59.93	0	100	0	12.50	55.71	100	100	56.35
Cannabis (vs. no cannabis)	100	28.10	20	42.86	100	36.43	1.19	100	100	0.21	3.08	73.81
Non-smoker (vs. smoker)	100	1.31	20	48.43	100	100	100	100	100	0.42	3.08	58.99
Non-drinker (vs. drinker)	100	28.10	100	100	100	36.43	1.19	100	100	0.84	100	59.26
PANSS Positive	100	30.72	100	39.72	100	62.50	0	100	56.43	100	3.08	100
PANSS Negative	100	30.07	100	41.81	100	62.50	0	100	56.43	100	3.08	100
PANSS General Psychopathology	100	30.72	100	42.51	100	62.50	0.60	100	56.43	100	3.08	100
PANSS Total Score	100	32.68	100	44.25	100	63.57	0.60	100	56.43	100	3.08	100
SAPS	100	74.51	100	100	3.51	100	100	100	100	0.21	100	0.26
SANS	100	75.16	100	100	3.51	100	100	100	100	0.84	100	0.53
GAF	100	71.24	100	59.58	1.75	45.00	13.10	46.88	55.71	100	100	100
BPRS Total Score	100	100	100	100	1.75	100	100	100	100	0.84	100	100
Family history of psychosis (vs. none)	32.17	100	100	100	100	100	100	100	100	100	100	49.74
Family history of mental health disorders (vs. none)	32.17	73.86	100	100	100	100	100	100	100	100	100	49.74
Highest education qualification	4.20	100	18.00	30.66	22.81	100	100	3.13	55.71	100	100	100
Ethnicity	2.80	82.35	0	0	0	3.57	1.19	100	100	0.63	100	100
Length of follow-up (years)	0	0	0	0	0	0	0	0	0	0	100	0

Table 17. Descriptive statistics stratified by outcome.

Variable	Outcome		
	Total	**Non-TR**	**TR**
Age of onset (years)	26.84 (9.20)	27.26 (9.40)	24.50 (7.60)
Age at baseline (years)	27.92 (9.17)	28.32 (9.32)	25.62 (7.88)
DUP (days)	298.50 (776.48)	299.81 (788.77)	290.29 (695.78)
Female (vs. male)*	792 (39.74)	698 (41.23)	94 (31.33)
Body Mass Index	24.01 (4.56)	23.93 (4.62)	24.57 (4.12)
Being in a relationship (vs. single)*	326 (20.01)	294 (21.75)	32 (11.55)
Ever being in a relationship (vs. never)*	447 (27.44)	405 (29.96)	42 (15.16)
Living alone (vs. with others)*	310 (26.63)	259 (26.62)	51 (26.70)
Living with non-parents (vs. with parents)*	315 (45.19)	269 (46.62)	46 (38.33)
Living with non-family (vs. with family)*	51 (6.48)	43 (6.48)	8 (6.50)
Living independent (vs. supported accommodation)*	298 (72.86)	268 (72.83)	30 (73.17)
Employed (vs. unemployed)*	243 (41.05)	208 (42.11)	35 (35.71)
Years in Education	12.66 (2.92)	12.79 (2.93)	11.85 (2.74)
Cannabis (vs. no cannabis)*	566 (43.74)	491 (43.30)	75 (46.88)
Non-smoker (vs. smoker)*	461 (44.76)	418 (46.09)	43 (34.96)
Non-drinker (vs. drinker)*	412 (38.18)	380 (39.26)	32 (28.83)
PANSS Positive	14.57 (5.80)	14.41 (5.78)	15.74 (5.88)
PANSS Negative	14.75 (5.98)	14.36 (5.74)	17.57 (6.90)
PANSS General Psychopathology	31.02 (7.88)	30.61 (7.65)	33.96 (8.82)
PANSS Total Score	60.88 (15.55)	59.85 (15.14)	68.34 (16.52)
SAPS	11.55 (3.98)	11.46 (3.96)	12.32 (4.06)
SANS	8.04 (6.56)	7.88 (6.47)	9.34 (7.20)
GAF	47.80 (16.79)	48.46 (17.17)	44.73 (14.59)
BPRS Total Score	63.12 (13.83)	62.73 (13.72)	65.06 (14.26)
Family history of psychosis (vs. none)*	116 (30.21)	98 (29.79)	18 (32.73)
Family history of mental health disorders (vs. none)*	175 (41.27)	151 (41.26)	24 (41.38)
No qualifications (vs. reference)*	117 (17.97)	86 (17.48)	31 (19.50)
Basic education qualification (vs. reference)*	190 (29.19)	135 (27.44)	55 (34.59)
Further education qualification (reference)*	230 (35.33)	171 (34.76)	59 (37.11)
Higher education qualification (vs. reference)*	114 (17.51)	100 (20.33)	14 (8.81)
Non-European (reference)*	1107 (68.89)	919 (69.57)	188 (65.73)
Black (vs. reference)*	322 (20.04)	249 (18.85)	73 (25.52)
Asian/Other (vs. reference)*	178 (11.08)	153 (11.58)	25 (8.74)
Length of follow-up (years)	4.13 (3.13)	3.89 (2.86)	5.37 (4.04)

NB: N (%) is reported for categorical variables and mean (standard deviation) for continuous variables. * indicates categorical variables where N (%) is shown for the category outside the brackets.

14.2. Appendix B

Psychological Medicine

cambridge.org/psm

Invited Review

Cite this article: Smart SE, Kępińska AP, Murray RM, MacCabe JH (2019). Predictors of treatment resistant schizophrenia: a systematic review of prospective observational studies. *Psychological Medicine* 1–10. https://doi.org/10.1017/S0033291719002083

Received: 10 January 2019
Revised: 24 June 2019
Accepted: 25 July 2019

Key words:
First episode; longitudinal; prediction; psychosis; schizophrenia; treatment resistant

Author for correspondence:
Sophie E Smart, E-mail: sophie.smart@kcl.ac.uk

CAMBRIDGE UNIVERSITY PRESS

Predictors of treatment resistant schizophrenia: a systematic review of prospective observational studies

S. E. Smart, A. P. Kępińska, R. M. Murray and J. H. MacCabe

Department of Psychosis Studies, Institute of Psychiatry, Psychology, & Neuroscience, King's College London, 16 de Crespigny Park, London, SE5 8AF, UK

Abstract

Treatment-resistant schizophrenia, affecting approximately 20–30% of patients with schizophrenia, has a high burden both for patients and healthcare services. There is a need to identify treatment resistance earlier in the course of the illness, in order that effective treatment, such as clozapine, can be offered promptly. We conducted a systemic literature review of prospective longitudinal studies with the aim of identifying predictors of treatment-resistant schizophrenia from the first episode. From the 545 results screened, we identified 12 published studies where data at the first episode was used to predict treatment resistance. Younger age of onset was the most consistent predictor of treatment resistance. We discuss the gaps in the literature and how future prediction models can identify predictors of treatment response more robustly.

Predictors of treatment-resistant schizophrenia: A systematic review of prospective observational studies

For approximately a third of patients with schizophrenia, standard antipsychotic medications do not adequately alleviate their psychotic symptoms (Conley and Kelly, 2001). This subgroup is termed treatment-resistant schizophrenia (TRS). The most common clinical and research criteria used for TRS is the failure to respond to two trials of non-clozapine antipsychotics, of adequate dose and duration (Suzuki *et al*., 2011; Howes *et al*., 2017).

Patients with TRS have higher rates of unemployment, worse quality of life, and poorer social and occupational functioning than people who respond to treatment (Iasevoli *et al*., 2016). Researchers have estimated that the direct healthcare costs for TRS in the US is 3–11-fold higher than for the schizophrenia population as a whole, with multiple hospitalisations accounting for a large proportion of this cost (Kennedy *et al*., 2014). In England, 25–50% of the National Health Service's (NHS) £11.8 billion mental health budget is allocated to schizophrenia services and TRS is thought to contribute a large proportion of these costs (Andrews *et al*., 2012; Killaspy *et al*., 2013).

Clozapine is the only antipsychotic recommended for TRS and is more effective than other antipsychotics in alleviating psychotic symptoms in patients with TRS (Kane *et al*., 1988; Siskind *et al*., 2016; Taylor, 2017). However, owing to its adverse effects, clozapine is only licenced in the UK (NICE, 2014) and most other developed countries (Warnez and Alessi-Severini, 2014) as a third-line treatment. Nevertheless, evidence suggests that TRS is often not recognised promptly, and that clozapine is offered after a delay of some years or not at all. According to treatment guidelines, the earliest that patients can be diagnosed with TRS, and prescribed clozapine, is 12 weeks after commencing antipsychotic treatment; however, Howes *et al*. (2012) report an average delay of 3.9 years, suggesting that there is considerable scope to shorten this period of inadequate treatment. Furthermore, patients with a shorter delay before clozapine initiation show a better symptomatic response to clozapine (Yoshimura *et al*., 2017).

Thus, there is a need to identify patients-who are likely to develop TRS-earlier in the course of their illness and expedite their access to specialist treatment; this may require moving beyond the current definition of TRS towards criteria based upon predictors and biomarkers, which quantify a patient's risk of developing TRS. If predictors of TRS can be identified, they may be useful in three ways: firstly, to identify TRS patients earlier in treatment so that they can be offered effective treatments earlier; secondly, to identify patients for clinical trials of interventions for TRS; and thirdly, to improve our understanding of the aetiology of TRS.

We present a comprehensive systematic review of all prospective observational studies in schizophrenia populations, which report baseline predictors of TRS. We focused solely on prospective observational studies to draw clearer conclusions regarding the causal relationship

between predictors and TRS in naturalistic settings over a long follow-up, and because only longitudinal studies can identify risk factors at first episode that might predict TRS.

Method

Inclusion/exclusion criteria

Studies were included if they met the following inclusion criteria: (1) participants were diagnosed with schizophrenia, schizophreniform disorder, schizoaffective disorder, and/or a psychotic disorder; we did not exclude studies that *also* included affective disorders or substance-induced psychosis, given the diagnostic uncertainly around the first episode of psychosis; (2) participants were followed from the first episode or first treatment with antipsychotics; (3) the majority of participants were aged between 16 and 64 at baseline (we excluded studies that focused exclusively on children or older adults); (4) data were collected prospectively from the first episode; (5) the outcome was a categorical definition of TRS, established using longitudinal prospective medication history; and (6) a non-TRS comparison group was recruited and followed up in the same manner as the TRS group. Studies were excluded if (1) they were clinical trials, or if non-antipsychotic treatments, such as CBT or ECT, were administered as part of the study procedure; (2) the study focussed exclusively on early or late-onset schizophrenia; or (3) inferential statistics measuring the association between baseline variables and TRS were not reported, and our subsequent requests to the authors for unpublished data were unsuccessful.

Defining TRS

Only recently has attention been given to the standardisation of TRS criteria (Farooq *et al*., 2013; Suzuki *et al*., 2012; Lee *et al*., 2015; Howes *et al*., 2017); therefore, we did not restrict studies to one definition of TRS. We did, however, only include studies with a categorical definition of TRS to capture the key underlying concept-at least two treatment failures – and differentiate TRS from relative measures of response/nonresponse. If patients took clozapine at follow-up, we inferred that they met criteria for TRS. Clozapine prescription is likely to underestimate the true proportion of patients with TRS (Howes *et al*., 2012), but it is a pragmatic criterion, since clozapine is only used for TRS, except in very rare indications (e.g. psychosis in the context of Parkinson's disease or for people who suffer severe side-effects to other antipsychotics).

Literature search

Studies were identified by searching Pubmed, PsychINFO (up to October 2017), Medline (up to October 2017), Embase (up to October 2017), and OpenGrey on the 1 November 2017. In addition, we examined the first 20 pages of Google Scholar using terms 'predictor AND treatment-resistant AND schizophrenia' on 3 January 2018. No restrictions were placed on the publication date, but searches were restricted to the titles and abstracts of papers (and subject headings in Medline, Embase, and PsychINFO), studies published in English, and studies using human participants. Search terms for Pubmed were as follows: '((treatment resistant) OR (treatment resistance) OR (treatment refractory)) AND (schizophrenia) AND ((longitudinal) OR (prospective))'. Search strategies for other databases can be found in Appendix 1. We screened the title and abstracts of all identified studies and then performed full-text screening of all potentially eligible studies. Potentially eligible studies were cross-referenced; additional relevant studies were identified by hand-searches of the references, and by screening papers which had previously cited these studies. Each additional paper was also hand-searched until no new studies were identified. When full-text articles were not available, the corresponding author was contacted. Author SES conducted the initial screening, with APK independently screening the studies identified through database searches and all studies identified through cross-referencing.

Quality assessment

We followed the PRISMA guidelines for reporting systematic reviews (Liberati *et al*., 2009).

Study quality was assessed using the Newcastle-Ottawa Scale (NOS) for cohort studies (http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp). Eight items measure the selection, comparability, and outcome of each study. These items were modified for this review, for example, follow-up needed to have been longer than one year to score on the item concerning adequate duration of follow-up (see Appendix 2). Authors SES and APK independently rated each study on the NOS (Appendix 3), any differences in rating were discussed between authors and final ratings were a consensus.

When available, we report adjusted hazard (HR) or odds ratios (OR) with 95% confidence intervals (95% CI) in parentheses, for predictors measured at baseline.

Results

A total of 12 studies were identified for inclusion in this review. Study screening is depicted in Fig. 1 and a summary of the number of participants recruited into each study is presented in Table 1. Database searches identified 545 records, 293 of which were duplicates and removed. A total of 252 records were screened and 248 were excluded. The main reasons for exclusion were: the study did not follow participants from the first episode or first treatment with antipsychotics (31%), participants recruited after TRS had been identified (29%), and an outcome other than TRS was reported (23%). The remaining four records were examined in more detail, as were the eight records identified by cross-referencing. Only duplicates were identified through Google Scholar. Of the 12 studies, 11 were published in peer-reviewed academic journals. One study was unpublished (Chan *et al*., 2014), however, after corresponding with the authors, a full report was identified on the funding body's website containing enough information to be included in this review (https://rfs1.fhb.gov.hk/app/fundedsearch/projectdetail.xhtml?id=1363).

Of the 12 included studies, eight presented original data and four presented data on additional exposures within the same cohort as a previous study, or a subset thereof. Of these eight, three were population cohort studies. Both Sorensen *et al*. (2014) and Wimberley *et al*. (2016b) used Danish population registers: data was extracted from multiple national databases and linked using a unique personal identification number. Additional analyses of Wimberley *et al*. (2016b)'s data tested whether urbanicity (Wimberley *et al*., 2016*a*), the polygenic risk score for schizophrenia (PRS-SZ; Wimberley *et al*., 2017), functioning (Horsdal *et al*., 2017*b*), and C-reactive protein levels (Horsdal *et al*., 2017*a*) could predict TRS. The third population

https://doi.org/10.1017/S0033291719002083

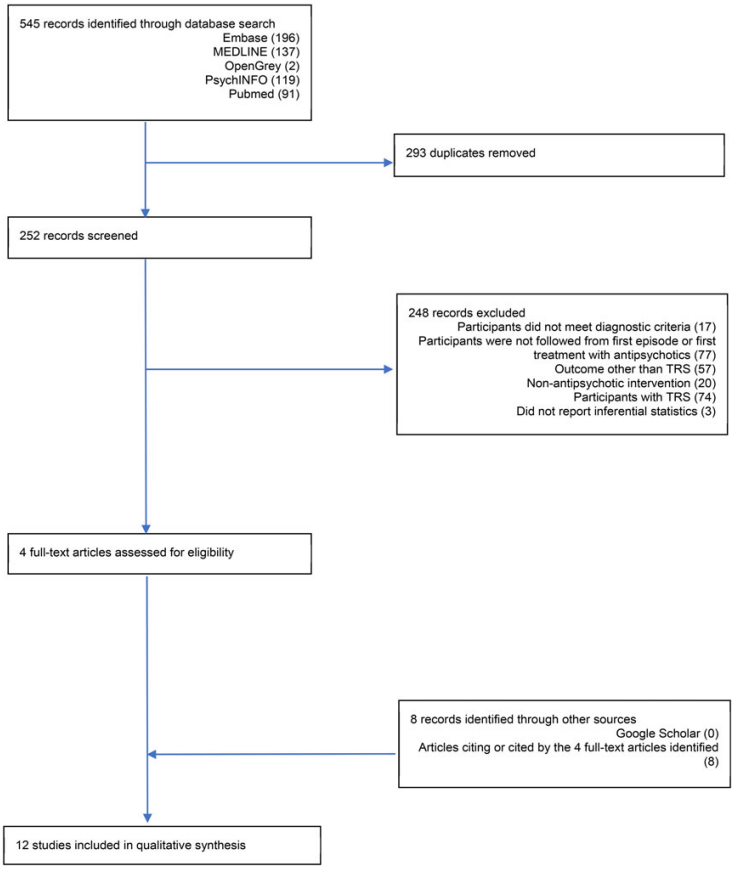


Fig. 1. PRISMA flow diagram.

cohort came from South Korea (Kim *et al.*, 2017). The remaining five studies analysed longitudinal first episode psychosis patient cohorts (Meltzer *et al.*, 1997; Chan *et al.*, 2014; Lally *et al.*, 2016; Üçok *et al.*, 2016; Demjaha *et al.*, 2017).

In population registries, a proxy definition of first-episode psychosis is required. In the Danish studies, the first International Classification of Diseases (ICD; World Health Organization, 1993) diagnosis of schizophrenia was used to define the baseline cohort. The South Korean study used ICD diagnosis of schizophrenia and the first use of antipsychotics to define the baseline cohort. When using diagnoses, the first episode is likely to be later in the disease course, when compared to cohort studies. Additional study characteristics, including information about recruitment, diagnoses, and criteria for TRS and non-TRS can be found in Appendix 4. The variables measured, and tested as predictors of TRS, varied considerably across studies, therefore this information is summarised in Table 2. Appendix 5 contains the unadjusted and adjusted OR/HR, when these were reported.

Predictors of TRS

Chan *et al.* (2014) analysed a subsample of a first episode cohort who presented to mental health services over a five-year period and used clozapine prescription as a definition of TRS. As this was a case-control study including all patients with TRS and a ratio of two non-TRS patients for every TRS patient, the

https://doi.org/10.1017/S0033291719002083

Table 1. The twelve studies included in this review, with details on the number of participants recruited and the length of follow-up

	Number of participants							
Study	Baseline	Follow-up	(%)	TRS	(%)	Non-TRS	(%)	Length of follow-up (years)
Chan *et al.* (2014)	469	469	(100)	160	(34)	309	(66)	16
Demjaha *et al.* (2017)	557	274	(49)	62	(23)	212	(77)	10
Horsdal *et al.* (2017*a*)*	390	390	(100)	52	(13)	338	(87)	2
Horsdal *et al.* (2017*b*)*	3252	3252	(100)	359	(11)	2893	(89)	2
Kim *et al.* (2017)	114 749	NR		NR		NR		NR
Lally *et al.* (2016)	283	240	(85)	81	(34)	159	(66)	5
Meltzer *et al.* (1997)	322	322	(100)	196	(61)	126	(39)	4
Sorensen *et al.* (2014)	5968	5328	(89)	1223	(23)	4105	(77)	34
Üçok *et al.* (2016)	187	105	(56)	28	(27)	77	(73)	2
Wimberley *et al.* (2016*a*)*	13 349	13 349	(100)	2313	(17)	11 036	(83)	17
Wimberley *et al.* (2016*b*)	9332	8044	(86)	1703	(21)	6341	(79)	14
Wimberley *et al.* (2017)*	862	862	(100)	181	(21)	681	(79)	11

NB: NR, not reported; *, analysis of data also presented in Wimberley *et al.* (2016*b*)

prevalence of TRS could not be calculated. The two groups were matched on baseline diagnosis. Chan *et al.* (2014) included age of onset, duration of untreated psychosis (DUP; days), duration of first episode, years of education, Premorbid Adjustment Scale (PAS) adult (19 + years) subscale score (Cannon-Spoor *et al.*, 1982), substance misuse history, and the number of relapses in the first three years, in a Cox proportional hazard regression. The model significantly predicted TRS (Chi-square = 66.11, df = 7, p = <0.0001). While number of relapses in the first three years significantly predicted TRS, the only baseline predictors significantly associated with TRS were younger age of onset (HR = 0.88, 95% CI = 0.83–0.94) and poorer premorbid functioning (indicated by higher scores) according to the PAS (HR = 3.22, 95% CI = 1.43–7.23).

Demjaha *et al.* (2017) analysed data from the AESOP study, which recruited first episode patients over a three-year period and followed them up ten years later. The researchers entered gender, diagnosis, age of onset, negative symptoms, mode of onset, DUP (weeks), and ethnicity into a multivariate penalised logistic regression. The model selected five variables that predicted TRS: a diagnosis of schizophrenia at baseline (instead of psychotic depression; OR = 0.41, or psychotic mania; OR = 0.52), younger age of onset (years, OR = 0.97), higher severity of negative symptoms (OR = 1.09), an insidious mode of onset (instead of acute; OR = 1.28), and longer DUP (OR = 1.0013). Goodness-of-fit was measured using McFadden's pseudo R^2 and correct classification rates were measured using the Brier score. A McFadden's pseudo R^2 between 0.20 and 0.40 is considered a good model fit. The Brier score is used to evaluate predictive models; if the incidence of TRS is 23%, as estimated from Demjaha *et al.* (2017), a Brier score of 0 would be a perfect model while a score of 0.177 would be a non-informative model (Steyerberg *et al.*, 2010). Demjaha *et al.* (2017) reported a McFadden's pseudo R^2 of 0.10 and a Brier score of 0.146, suggesting that their model is not a good fit of the data nor is it a good classifier of TRS.

Kim *et al.* (2017), in their South Korean population cohort, estimated the cumulative incidence of clozapine use using the Kaplan–Meier method and log-rank test. They reported that younger age of onset predicted TRS. Unlike Chan *et al.* (2014) and Demjaha *et al.* (2017), Kim *et al.* (2017) examined age of onset categorically: defining younger age of onset as those aged between 15–20 years of age, and comparing them to a middle-onset group (21–44 years of age) and a late-onset group (45–64 years of age). Kim *et al.* (2017) also found, using the Walter–Elwood method (Walter and Elwood, 1975), a higher incidence of clozapine use, in those born during winter (December to February) when compared to those born in summer (June to August). This pattern remained true when stratifying season of birth by age of onset. Kim *et al.* (2017) reported no measures of overall model fit.

Lally *et al.* (2016) recruited first episode patients over a five-year period and used electronic medical records to follow them up five years later. They entered age of onset, Positive and Negative Symptom Scale (PANSS; Kay *et al.*, 1987) scores, Global Assessment of Functioning (GAF; Hall, 1995) disability score, and GAF symptom scores into a penalised logistic regression, controlling for living arrangements, employment status, and alcohol/substance misuse during the follow-up period. Lally *et al.* (2016) included the PANSS total score, the positive, negative and general psychopathology subscale scores, as well as two individual items: lack of insight and conceptual disorganisation. None of the PANSS or GAF variables predicted TRS. Age at first contact with mental health services was split into four categories: 18–20, 21–25, 26–30, >31 years. Only age of onset between 18 and 20 years, compared to all other age groups, significantly predicted TRS (OR = 2.49, 95% CI = 1.25–4.94). The authors did not report the overall model fit. Age of onset was subsequently stratified by gender and ethnicity. Age of onset, between 18 and 20, only predicted TRS in males (OR = 2.13, 95% CI = 1.35–7.23) or those of black ethnicity (OR = 3.71, 95% CI = 1.44–9.56).

Meltzer *et al.* (1997) recruited patients at first admission to hospital for schizophrenia or schizoaffective disorder and followed them up for approximately four years. The authors examined the age of onset and gender in relation to TRS using a two-way analysis of variance (ANOVA). Gender was not associated with TRS but younger age of onset was. As males had a

https://doi.org/10.1017/S0033291719002083

Table 2. The variables which have been tested as predictors of TRS in the twelve studies included in this review

	Chan *et al.* (2014)	Demjaha *et al.* (2017)	Kim *et al.* (2017)	Lally *et al.* (2016)	Meltzer *et al.* (1997)	Sorensen *et al.* (2014)	Üçok *et al.* (2016)	Wimberley *et al.* (2016*b*)
Younger age of onset								
Alcohol misuse during follow-up period								
Antipsychotic polypharmacy during follow-up period								
C-reactive protein								
Comorbid diagnosis of personality disorder								
Comorbid diagnosis of suicide attempts								
Schizophrenia diagnosis								
Paranoid schizophrenia diagnosis								
Duration of first episode								
Longer DUP								
Early parental loss								
Lower education qualification								
Fewer years in education								
Employment status								
Black ethnicity								
Family history of schizophrenia								
Worse functioning								
Worse premorbid functioning								
Male								
Living arrangements								
Living arrangements during follow-up period								
Marital status								
Mode of onset								
Paternal age								
Polygenic risk score for schizophrenia								
Relapse despite adherence								
Relapse in first 6 months								
More relapses in the first three years								
Born in Autumn/winter								
Substance misuse								

(*Continued*)

https://doi.org/10.1017/S0033291719002083

Table 2. (*Continued.*)

	Chan *et al.* (2014)	Demjaha *et al.* (2017)	Kim *et al.* (2017)	Lally *et al.* (2016)	Meltzer *et al.* (1997)	Sorensen *et al.* (2014)	Üçok *et al.* (2016)	Wimberley *et al.* (2016*b*)
Substance misuse during follow-up period								
More symptoms of psychosis		■		■				
Inpatient								■
Urbanicity at birth								
Living in a more rural area at diagnosis								■
Violent offence								
	Chan *et al.* (2014)	Demjaha *et al.* (2017)	Kim *et al.* (2017)	Lally *et al.* (2016)	Meltzer *et al.* (1997)	Sorensen *et al.* (2014)	Üçok *et al.* (2016)	Wimberley *et al.* (2016*b*)

NB: grey squares, variables not significantly associated with TRS; dark grey squares, variables significantly associated with TRS; all analyses using the Wimberley *et al.*, Danish dataset were grouped under Wimberley *et al.* (2016*b*)

younger age of onset than females, the researchers examined the associations between age of onset and gender in more depth using simple effects ANOVA. In the non-TRS group, males had a younger age of onset ($F = 6.6$, df = 1, $p < 0.01$), however, in the TRS group, there was no difference in age of onset between males and females. Meltzer *et al.* (1997) calculated the conditional probability of a patient having TRS given their age of onset. For those aged between 15 and 18 years old, the probability of developing TRS was between 32% and 38% for both males and females.

Sorensen *et al.* (2014), in their Danish population cohort, entered a season of birth into a Cox proportion hazard regression adjusted for birth year and gender. The model did not significantly predict TRS. However, the authors found that being born in autumn (September to November), compared to spring (March to May), predicted TRS (HR = 1.24, 95% CI = 1.06–1.46). Unlike in Kim *et al.* (2017)'s study, being born in winter (December to February) failed to predict TRS.

Üçok *et al.* (2016) analysed a subsample of patients recruited into an ongoing first episode schizophrenia study. Üçok *et al.* (2016) entered the age of onset, DUP (days), first relapse despite adherence to antipsychotic treatment, relapse in the first six months, and antipsychotic polypharmacy during follow-up, into logistic regression. The authors did not report the overall model fit. Only first relapse despite adherence to antipsychotic treatment and antipsychotic polypharmacy predicted TRS. No baseline variables predicted TRS.

Wimberley *et al.* (2016*b*), in their Danish population cohort, entered twenty-three variables into a Cox proportion hazard regression. These variables included: gender, age at first schizophrenia diagnosis as a proxy for age of onset, family history of schizophrenia in first-degree relatives, winter birth (December to March), paternal age, parental loss before the age of 18, living alone, conviction for a violent offence before first schizophrenia diagnosis, level of education, employment status, urbanicity at first schizophrenia diagnosis, admission to psychiatric hospital before first schizophrenia diagnosis, schizophrenia subtype (paranoid *v.* all others), comorbid psychiatric diagnosis before first schizophrenia diagnosis, antipsychotic prescription in the year before first schizophrenia diagnosis, antidepressant prescription in the year before first schizophrenia diagnosis, and benzodiazepine prescription in the year before first schizophrenia diagnosis. Goodness-of-fit was measured using McFadden's pseudo R^2 and correct classification rates using Harrell's C statistic; a C statistic of 0.5 would be a non-informative model while a score of 1 would be a perfect model. Wimberley *et al.* (2016*b*) report a McFadden's pseudo R^2 of 0.027 and a Harrell's C statistic of 0.70, suggesting that this model is a good fit of the data and reasonable classifier of TRS. At baseline, younger age of onset (years, HR = 0.96, 95% CI = 0.95–0.97), living in less urban areas (rural *v.* capital area, HR = 1.44, 95% CI = 1.25–1.65), higher education (higher *v.* primary education, HR = 0.88, 95% CI = 0.79–0.98), psychiatric hospital admission at diagnosis (HR = 2.07, 95% CI = 1.87–2.29), having spent more than 30 bed-days in a psychiatric hospital in the year before diagnosis (HR = 1.54, 95% CI = 1.35–1.75), paranoid subtype diagnosis (HR = 1.24, 95% CI = 1.13–1.37), comorbid personality disorder (HR = 1.24, 95% CI = 1.11–1.39), comorbid suicide attempt (HR = 1.21, 95% CI = 1.07–1.39), antipsychotic use (HR = 1.51, 95% CI = 1.35–1.69), antidepressant use (HR = 1.15, 95% CI = 1.03–1.29), and benzodiazepines use (HR = 1.22, 95% CI = 1.10–1.37), all predicted TRS. Data on additional exposures within the same cohort, or a subset thereof, were published separately. Lower levels of urbanicity (Wimberley *et al.*, 2016*a*) and severely impaired functioning (a GAF functioning score ⩽30; Horsdal *et al.*, 2017*b*) predicted TRS, but the polygenic risk score for schizophrenia (PRS-SZ; Wimberley *et al.*, 2017) and C-reactive protein levels (Horsdal *et al.*, 2017*a*) did not predict TRS.

Subcategories of TRS

Some patients have little or no response to antipsychotic treatment from the onset of their illness, while others initially respond to medication and then later develop TRS. Two of the studies in our review reported comparisons between subgroups TRS patients; early-onset TRS was operationalised as meeting criteria for TRS from the onset of schizophrenia and delayed-onset TRS as meeting criteria after a period of symptomatic remission. Chan *et al.* (2014) found no differences, in demographics, clinical characteristics, or premorbid functioning, between early-onset

https://doi.org/10.1017/S0033291719002083

TRS ($N = 17$, 11.64%) and delayed-onset TRS ($N = 129$, 88.36%). Lally *et al.* (2016) found no differences in demographics between the two groups, but the early-onset TRS group ($N = 56$, 70%) had a younger mean age of onset than the delayed-onset TRS group ($N = 24$, 30%).

Discussion

This review identified twelve research papers that examined predictors of TRS. Seven of the studies included in this review tested the age of onset as a predictor, and six reported that younger age of onset predicted TRS. Given that multiple definitions of the age of onset – age of onset of psychotic symptoms, age of first diagnosis of schizophrenia, age of first contact with mental health services – were reported and data was treated both continuously and categorically, this is a robust finding. Other potential risk factors, that have been identified by more than one study, include diagnosis, level of functioning, male gender, and season of birth.

A recent meta-analysis linked younger age of onset to multiple poor outcomes in schizophrenia: more hospitalisations, more negative symptoms, more relapses, poorer social/occupational functioning, and poorer global outcome (Immonen *et al.*, 2017). Many of these poor outcomes have also been associated with TRS. Immonen *et al.* (2017) found that males had a younger age of onset and, therefore, samples with a higher proportion of males tended to show stronger associations between age of onset and outcomes. In the studies included in this review, the association between age of onset and TRS is unlikely to be wholly confounded by gender, as the proportion of males ranged from 49% (Kim *et al.*, 2017) to 67% (Lally *et al.*, 2016) and the studies which controlled for gender still showed an effect of age of onset (Meltzer *et al.*, 1997; Lally *et al.*, 2016; Wimberley *et al.*, 2016*b*; Demjaha *et al.*, 2017). In schizophrenia, age of onset has been thought to reflect genetic liability for the disease; younger age of onset has been associated with an increased familial risk of schizophrenia (Hilker *et al.*, 2017; Byrne *et al.*, 2018). Could, therefore, TRS be the result of increased genetic risk? While Wimberley *et al.* (2017) found no association between PRS-SZ and TRS, other work published by Frank *et al.* (2014) reports that an increased PRS-SZ is associated with TRS. In addition, rare copy number variations have been associated with both TRS (Martin and Mowry, 2015) and childhood-onset schizophrenia (Addington and Rapoport, 2009). Therefore, patients with TRS, who also have a younger age of onset, may have a more salient genetic influence than later-onset cases, although further work is required to substantiate this claim.

This review complements previous reviews by Gillespie *et al.* (2017) and Carbon and Correll (2014). Gillespie *et al.* (2017) examined studies comparing patients with treatment-resistant to patients with treatment-responsive schizophrenia. They included all study methodologies, but excluded studies where treatment-responsiveness was defined solely as not meeting treatment-resistant criteria. Carbon and Correll (2014) examined studies identifying predictors of response and remission. The researchers focused on first-episode psychosis studies where participants were followed up for five years. Some of the predictors of TRS, identified in this review, were found to be associated with less chance of response/remission by Carbon and Correll (2014), e.g. younger age of illness onset, poor premorbid adjustment, being male, lower level of education, living in a rural environment, diagnosis of schizophrenia, longer duration of untreated psychosis, poorer functioning, and worse psychopathology. However, Carbon and Correll (2014) also associated less chance of response/remission with being single, family history of psychosis, greater cognitive dysfunction, more family conflicts, and substance misuse; characteristics not identified as predictors of TRS. There was relatively little overlap between this review and Gillespie *et al.* (2017)'s review. In terms of studies included, only Meltzer *et al.* (1997)'s study was included in both reviews. In terms of characteristics associated with TRS, Gillespie *et al.* (2017) identified five neuroimaging studies, nine gene-association studies, and two studies of neurocognitive function, and these studies were not included in our review. The examination of biological markers, associated with TRS, within longitudinal study designs is rare; this is understandable for genome-wide association studies, which require large sample sizes more easily acquired using a cross-sectional methodology. However, there is a clear gap in the literature investigating biological markers that change over time (for example, proinflammatory cytokines or differently methylated positions within the epigenome) and TRS as an outcome. In terms of neuroimaging research, a review by Nakajima *et al.* (2015) found only five studies which compared patients with TRS to non-TRS patients, none of which had identified neural correlates of TRS. McGuire and Dazzan (2017) highlight only one study where neuroimaging data predicted a six-year, non-remitting course of illness. Longitudinal imaging studies of TRS are still relatively rare and constitute another gap in the literature.

Of the studies included in this review, few identified characteristics of abnormal neurodevelopment as predictors of TRS, despite neurodevelopment changes being linked with schizophrenia. The neurodevelopmental theory of schizophrenia proposes that disrupted normal development, in utero or early infancy, leads to deficits in psychophysiological and neurological functioning in childhood or early adolescence, and eventually to prodromal or diagnostic symptoms of schizophrenia (Jablensky *et al.*, 2017; Murray *et al.*, 2017). Previous research has linked characteristics of abnormal development with TRS; higher rates of minor physical anomalies (Lin *et al.*, 2015), more neurological soft signs (de Bartolomeis *et al.*, 2018), poor verbal intelligence and fluency (Kravariti *et al.*, 2018), and poor verbal memory (Joober *et al.*, 2002; de Bartolomeis *et al.*, 2013). None of the studies in this review included variables measuring physiology during development or cognition at the first episode. Only Chan *et al.* (2014) examined premorbid functioning, retrospectively using the PAS. They found no difference, between the TRS and non-TRS groups, in terms of functioning during childhood, early adolescence, or late adolescent. There was a difference in functioning after the age of 19 and subsequently, worse functioning predicted TRS in their final model. If educational attainment can be considered a proxy for development only lower level of education qualification was found to significantly predict TRS (Wimberley *et al.*, 2016*b*). The number of years in education was not predictive of TRS (Chan *et al.*, 2014). Abnormal neurodevelopment and neuropsychology have not been sufficiently investigated as potential predictors of TRS.

Our review has illuminated some gaps in the existing literature, where potential predictors have not been fully investigated, however, we believe our review has captured all published work and identified predictors that, with further study, may prove to be clinically useful in determining treatment for patients with schizophrenia.

Strengths and limitations

The main strength of this review is that we have focused solely on studies that included temporal forecasting (observations at baseline that are used to predict outcomes at follow-up), and as

https://doi.org/10.1017/S0033291719002083

such eliminated recall bias and established a key component necessary for predictive models. All the studies included in this review are likely to be sufficiently powered to detect predictors of TRS. All the studies reported large sample sizes, and most followed participants for more than one year. Although no studies reported *a priori* power analysis, and only Meltzer *et al.* (1997) reported an *ad hoc* power analysis, we believe lack of power is unlikely to explain these results.

When attrition reduces the sample size at follow-up of longitudinal studies, consequently, statistical power is also reduced. For the studies we have reviewed, that reported on participants lost to follow-up, it is unlikely that the low attrition rates introduced bias. In particular, many studies used Cox proportional hazard regression; an analytic method that not only takes into account that individuals lost to follow-up may develop TRS, but also that individuals may develop TRS after the study endpoint. However, TRS, in particular, may be biased by attrition. There is a case both that TRS patients may be more likely to drop out of research studies due to their higher severity of symptoms and worse social and occupational functioning, and that responders are more likely to drop out as they lose touch with clinical services, but we are not aware of any published studies examining attrition in relation to treatment response.

One limitation to consider, when discussing the findings from these studies, is that some patients may have been misclassified. None of the studies included in this review explicitly accounted for adherence to medication, therefore characteristics may be predicting nonadherence rather than treatment resistance. None of the studies measured antipsychotic plasma levels, therefore characteristics may be predicting sub-therapeutic drug plasma levels, as a consequence of nonadherence, noncompliance, or pharmacokinetics, rather than treatment resistance. McCutcheon *et al.* (2015) found that 44% of patients referred to an outpatient service for clozapine treatment had sub-therapeutic conventional-antipsychotic plasma levels. On the other hand, it is unlikely that TRS patients have been wrongly classified as responders because the long follow-up periods allow plenty of time for a diagnosis of TRS to be established. Most studies had follow-ups longer than four years; the average delay before being treated for TRS estimated by Howes *et al.* (2012). The definitions of TRS, used in these studies, are pragmatic criteria: any predictors identified by these naturalistic studies are generalisable to real-world, clinical settings where adherence, compliance, or drug plasma levels influence treatment.

The use of multiple definitions of TRS is a problem across all TRS literature; Suzuki *et al.* (2011) reviewed 33 studies of prospective studies of pharmacological interventions for TRS and found that all 33 definitions of TRS were different. Howes *et al.* (2017) reviewed 42 clinical trials and found only two studies which used identical criteria. In addition, some studies use clozapine prescription as a proxy for TRS. When clozapine is underprescribed, supposed predictors of TRS may, in fact, represent predictors of clozapine initiation (e.g. clinicians' attitudes towards clozapine prescription). All of the studies, identified in this review, used existing data, not designed to examine TRS, and researchers had to established proxy definitions based on the data available to them. When evidence concerning predictors of TRS is not consistent, it can be hard to draw a clear conclusion about the validity of the predictor, yet when the evidence is consistent across studies, with different definitions, the predictor in question is highly likely to generalise to other cohorts and have clinical validity.

Finally, we must consider the statistical methodology used to establish predictors. A common misconception is that predictive accuracy can be inferred from explanatory accuracy. However, the two are different and should be assessed separately (Shmueli, 2010). Only three studies included in this review reported the overall model fit, and only two reported statistics that measure the predictive validity of the model. Additionally, in predictive modelling, variable selection and overfitting must be considered. Lally *et al.* (2016) and Demjaha *et al.* (2017) attempted to reduce overfitting by penalising regression coefficients. However, none of the studies used holdout data (training data), cross-validation, or external validation to evaluate the predictive power of models; the latter being the current 'gold-standard' approach. In terms of variable selection, the only methods reported were LASSO regression (Demjaha *et al.*, 2017) and step-wise selection using statistical significance (Chan *et al.*, 2014; Üçok *et al.*, 2016). Stepwise methods are no longer considered appropriate for explanatory models, but stepwise-type algorithms are very useful in predictive modelling (Shmueli, 2010), as long as the selection criteria rely on predictive power (e.g. Akaike information criterion) rather than explanatory power (e.g. statistical significance), as was the case in these studies. These methodological limitations must be taken into consideration when evaluating predictive models. The studies included in the review, on the whole, report analyses designed to identify explanatory variables of TRS. Future studies will need to use more robust prediction methods before moving from statistical prediction to clinical prediction.

Conclusion

The aim of this systemic literature review was to identify predictors of treatment-resistant schizophrenia from prospective longitudinal studies. In choosing to focus exclusively on longitudinal studies, we have filled a gap in the existing literature, and hope that consolidating this information will be of use to researchers attempting to identify clinical predictors of TRS and the biological mechanisms causing TRS. We have identified earlier age of schizophrenia-onset as a robust predictor of TRS, with evidence that male gender, autumn/winter birth, poor premorbid functioning and rural upbringing may also contribute. We have also highlighted gaps in the literature namely, studies examining neuroimaging, immune, and genetic markers of TRS. Examination of biological markers, particularly within the framework of a prospective longitudinal study, has the potential to go beyond simple prediction and add to our understanding of the underlying causes of TRS. In conclusion, while early identification of TRS is clinically important, we currently have very limited knowledge of its predictors.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291719002083.

Funding. This work was supported by the Medical Research Council (MRC) (S.E.S., R.M.M., & J.H.M., grant number: MR/L011794/1) and the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) South London at King's College Hospital NHS Foundation Trust South London (S.E.S.). A.P.K. was supported by the NIHR Biomedical Research Centre (BRC) at South London and Maudsley NHS Foundation Trust and King's College London.

The views expressed in this article are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care, the BRC, the MRC, or King's College London.

Conflict of interest. S.E.S. has received travel and accommodation expenses from H Lundbeck A/S. A.P.K. declares no conflicts of interest. R.M.M. has received honoraria for lectures from Janssen, Otsuka, Lundbeck and

Sunovian. J.H.M. has received travel and accommodation expenses and research funding from H Lundbeck A/S.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

Addington AM and Rapoport JL (2009) The genetics of childhood-onset schizophrenia: when madness strikes the prepubescent. *Current Psychiatry Reports* **11**, 156–161.

Andrews A, Knapp M, McCrone P, Parsonage M. and Trachtenberg M (2012) *Effective Interventions in Schizophrenia: the Economic Case*. A Report Prepared for the Schizophrenia Commission. Rethink Mental Illness: London.

Byrne M, Agerbo E and Mortensen PB (2018) Family history of psychiatric disorders and age at first contact in schizophrenia: an epidemiological study. *The British Journal of Psychiatry* **181**, s19–s25.

Cannon-Spoor HE, Potkin SG and Wyatt RJ (1982) Measurement of premorbid adjustment in chronic schizophrenia. *Schizophrenia Bulletin* **8**, 470–484.

Carbon M and Correll CU (2014) Clinical predictors of therapeutic response to antipsychotics in schizophrenia. *Dialogues in Clinical Neuroscience* **16**, 505–524.

Chan SKW, Chang WC, Lee EHM, Wong GHY, Hui CLM, Chen EYH and Sham PC (2014) Predictors of treatment resistant schizophrenia-spectrum disorder: 10-year retrospective study of first-episode psychosis. *Early Intervention in Psychiatry* **8**, 78.

Conley RR and Kelly DL (2001) Management of treatment resistance in schizophrenia. *Biological Psychiatry* **50**, 898–911.

de Bartolomeis A, Balletta R, Giordano S, Buonaguro EF, Latte G. and Iasevoli F (2013) Differential cognitive performances between schizophrenic responders and non-responders to antipsychotics: correlation with course of the illness, psychopathology, attitude to the treatment and antipsychotics doses. *Psychiatry Research* **210**, 387–395.

de Bartolomeis A, Prinzivalli E, Callovini G, D'Ambrosio L, Altavilla B, Avagliano C and Iasevoli F (2018) Treatment resistant schizophrenia and neurological soft signs may converge on the same pathology: evidence from explanatory analysis on clinical, psychopathological, and cognitive variables. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* **81**, 356–366.

Demjaha A, Lappin JM, Stahl D, Patel MX, MacCabe JH, Howes OD, Heslin M, Reininghaus UA, Donoghue K, Lomas B, Charalambides M, Onyejiaka A, Fearon P, Jones P, Doody G, Morgan C, Dazzan P and Murray RM (2017) Antipsychotic treatment resistance in first-episode psychosis: prevalence, subtypes and predictors. *Psychological Medicine* **47**, 1981–1989.

Farooq S, Agid O, Foussias G and Remington G (2013) Using treatment response to subtype schizophrenia: proposal for a new paradigm in classification. *Schizophrenia Bulletin* **39**, 1169–1172.

Frank J, Lang M, Witt SH, Strohmaier J, Rujescu D, Cichon S, Degenhardt F, Nöthen MM, Collier DA, Ripke S, Naber D and Rietschel M (2014) Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients. *Molecular Psychiatry* **20**, 150.

Gillespie AL, Samanaite R, Mill J, Egerton A and MacCabe JH (2017) Is treatment-resistant schizophrenia categorically distinct from treatment-responsive schizophrenia? A systematic review. *BMC Psychiatry* **17**, 12.

Hall RC (1995) Global assessment of functioning: a modified scale. *Psychosomatics* **36**, 267–275.

Hilker R, Helenius D, Fagerlund B, Skytthe A, Christensen K, Werge TM, Nordentoft M and Glenthøj B (2017) Is an early Age at illness onset in schizophrenia associated with increased genetic susceptibility? Analysis of data from the nationwide Danish Twin register. *EBioMedicine* **18**, 320–326.

Horsdal HT, Wimberley T, Benros ME and Gasse C (2017*a*) C-reactive protein levels and treatment resistance in schizophrenia—A Danish population-based cohort study. *Human Psychopharmacology: Clinical and Experimental* **32**, e2632-n/a.

Horsdal HT, Wimberley T, Köhler-Forsberg O, Baandrup L and Gasse C (2017*b*) Association between global functioning at first schizophrenia diagnosis and treatment resistance. *Early Intervention in Psychiatry* **12**, 1198–1202.

Howes OD, Vergunst F, Gee S, McGuire P, Kapur S and Taylor D (2012) Adherence to treatment guidelines in clinical practice: study of antipsychotic treatment prior to clozapine initiation. *The British Journal of Psychiatry* **201**, 481–485.

Howes OD, McCutcheon R, Agid O, de Bartolomeis A, van Beveren NJ, Birnbaum ML, Bloomfield MA, Bressan RA, Buchanan RW, Carpenter WT, Castle DJ, Citrome L, Daskalakis ZJ, Davidson M, Drake RJ, Dursun S, Ebdrup BH, Elkis H, Falkai P, Fleischacker W, Gadelha A, Gaughran F, Glenthoj BY, Graff-Guerrero A, Hallak JE, Honer WG, Kennedy J, Kinon BJ, Lawrie SM, Lee J, Leweke F, MacCabe JH, McNabb CB, Meltzer H, Moller H-J, Nakajima S, Pantelis C, Marques TR, Remington G, Rossell SL, Russell BR, Siu CO, Suzuki T, Sommer IE, Taylor D, Thomas N, Ucok A, Umbricht D, Walters JT, Kane J and Correll CU (2017) Treatment-resistant schizophrenia: treatment response and resistance in psychosis (TRRIP) working group consensus guidelines on diagnosis and terminology. *The American Journal of Psychiatry* **174**, 216–229.

Iasevoli F, Giordano S, Balletta R, Latte G, Formato MV, Prinzivalli E, De Berardis D, Tomasetti C and de Bartolomeis A (2016) Treatment resistant schizophrenia is associated with the worst community functioning among severely-ill highly-disabling psychiatric conditions and is the most relevant predictor of poorer achievements in functional milestones. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* **65**, 34–48.

Immonen J, Jääskeläinen E, Korpela H and Miettunen J (2017) Age at onset and the outcomes of schizophrenia: a systematic review and meta-analysis. *Early Intervention in Psychiatry* **11**, 453–460.

Jablensky A, McNeil TF and Morgan VA (2017) Barbara fish and a short history of the neurodevelopmental hypothesis of schizophrenia. *Schizophrenia Bulletin* **43**, 1158–1163.

Joober R, Rouleau GA, Lal S, Dixon M, O'Driscoll G, Palmour R, Annable L, Bloom D, Lalonde P, Labelle A and Benkelfat C (2002) Neuropsychological impairments in neuroleptic-responder vs. -nonresponder schizophrenic patients and healthy volunteers. *Schizophrenia Research* **53**, 229–238.

Kane J, Honigfeld G, Singer J and Meltzer H (1988) Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Archives of General Psychiatry* **45**, 789–796.

Kay SR, Fiszbein A and Opler LA (1987) The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* **13**, 261–276.

Kennedy JL, Altar CA, Taylor DL, Degtiar I and Hornberger JC (2014) The social and economic burden of treatment-resistant schizophrenia: a systematic literature review. *International Clinical Psychopharmacology* **29**, 63–76.

Killaspy H, Marston L, Omar RZ, Green N, Harrison I, Lean M, Holloway F, Craig T, Leavey G and King M (2013) Service quality and clinical outcomes: an example from mental health rehabilitation services in England. *The British Journal of Psychiatry* **202**, 28–34.

Kim JS, Park CM, Choi JA, Park E, Tchoe HJ, Choi M, Suh JK, Kim YH, Won SH, Chung YC, Bae KY, Lee SK, Park SC and Lee SH (2017) The association between season of birth, age at onset, and clozapine use in schizophrenia. *Acta Psychiatrica Scandinavica* **136**, 445–454.

Kravariti E, Demjaha A, Zanelli J, Ibrahim F, Wise C, MacCabe JH, Reichenberg A, Pilecka I, Morgan K, Fearon P, Morgan C, Doody GA, Donoghue K, Jones PB, Kacar AS, Dazzan P, Lappin J and Murray RM (2018) Neuropsychological function at first episode in treatment-resistant psychosis: findings from the AESOP-10 study. *Psychological Medicine* 1–11. https://doi.org/10.1017/ S0033291718002957

Lally J, Ajnakina O, Di Forti M, Trotta A, Demjaha A, Kolliakou A, Mondelli V, Marques T, Pariante C, Dazzan P, Shergil S, Howes O, David A, MacCabe J, Gaughran F and Murray R (2016) Two distinct patterns of treatment resistance: clinical predictors of treatment resistance in first-episode schizophrenia spectrum psychoses. *Psychological Medicine* **46**, 3231–3240.

Lee J, Takeuchi H, Fervaha G, Sin GL, Foussias G, Agid O, Farooq S and Remington G (2015) Subtyping schizophrenia by treatment response: antipsychotic development and the central role of positive symptoms. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie* **60**, 515–522.

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J and Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine* **6**, e1000100.

Lin AS, Chang SS, Lin SH, Peng YC, Hwu HG and Chen WJ (2015) Minor physical anomalies and craniofacial measures in patients with treatment-resistant schizophrenia. *Psychological Medicine* **45**, 1839–1850.

Martin AK and Mowry B (2015) Increased rare duplication burden genome-wide in patients with treatment-resistant schizophrenia. *Psychological Medicine* **46**, 469–476.

McCutcheon R, Beck K, Bloomfield MA, Marques TR, Rogdaki M and Howes OD (2015) Treatment resistant or resistant to treatment? Antipsychotic plasma levels in patients with poorly controlled psychotic symptoms. *Journal of Psychopharmacology* **29**, 892–897.

McGuire P and Dazzan P (2017) Does neuroimaging have a role in predicting outcomes in psychosis? *World Psychiatry* **16**, 209–210.

Meltzer HY, Rabinowitz J, Lee MA, Cola PA, Ranjan R, Findling RL and Thompson PA (1997) Age at onset and gender of schizophrenic patients in relation to neuroleptic resistance. *American Journal of Psychiatry* **154**, 475–482.

Murray RM, Bhavsar V, Tripoli G. and Howes O (2017) 30 years on: how the neurodevelopmental hypothesis of schizophrenia morphed into the developmental risk factor model of psychosis. *Schizophrenia Bulletin* **43**, 1190–1196.

Nakajima S, Takeuchi H, Plitman E, Fervaha G, Gerretsen P, Caravaggio F, Chung JK, Iwata Y, Remington G and Graff-Guerrero A (2015) Neuroimaging findings in treatment-resistant schizophrenia: a systematic review: lack of neuroimaging correlates of treatment-resistant schizophrenia. *Schizophrenia Research* **164**, 164–175.

NICE (2014) Psychosis and Schizophrenia in Adults: Treatment and Management: Updated Edition 2014. In *National Institute for Health and Clinical Excellence: Guidance* National Collaborating Centre for Mental Health: London.

Shmueli G (2010) To explain or to predict? *Statistical Science* **25**, 289–310.

Siskind D, McCartney L, Goldschlager R and Kisely S (2016) Clozapine v. first- and second-generation antipsychotics in treatment-refractory schizophrenia: systematic review and meta-analysis. *British Journal of Psychiatry* **209**, 385–392.

Sorensen HJ, Foldager L, Roge R, Pristed SG, Andreasen JT and Nielsen J (2014) An association between autumn birth and clozapine treatment in patients with schizophrenia: a population-based analysis. *Nordic Journal of Psychiatry* **68**, 428–432.

Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ and Kattan MW (2010) Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology* **21**, 128–138.

Suzuki T, Remington G, Mulsant BH, Rajji TK, Uchida H, Graff-Guerrero A and Mamo DC (2011) Treatment resistant schizophrenia and response to antipsychotics: a review. *Schizophrenia Research* **133**, 54–62.

Suzuki T, Remington G, Mulsant BH, Uchida H, Rajji TK, Graff-Guerrero A, Mimura M and Mamo DC (2012) Defining treatment-resistant schizophrenia and response to antipsychotics: a review and recommendation. *Psychiatry Research* **197**, 1–6.

Taylor DM (2017) Clozapine for treatment-resistant schizophrenia: still the gold standard? *CNS Drugs* **31**, 177–180.

Üçok A, Çıkrıkçılı U, Ergül C, Tabak Ö, Salaj A, Karabulut S and Correll CU (2016) Correlates of clozapine use after a first episode of schizophrenia: results from a long-term prospective study. *CNS Drugs* **30**, 997–1006.

Walter S and Elwood J (1975) A test for seasonality of events with a variable population at risk. *Journal of Epidemiology & Community Health* **29**, 18–21.

Warnez S. and Alessi-Severini S. (2014) Clozapine: a review of clinical practice guidelines and prescribing trends. *BMC Psychiatry* **14**, 102–102.

Wimberley T, Pedersen CB, MacCabe JH, Støvring H, Astrup A, Sørensen HJ, Horsdal HT, Mortensen PB and Gasse C (2016*a*) Inverse association between urbanicity and treatment resistance in schizophrenia. *Schizophrenia Research* **174**, 150–155.

Wimberley T, Støvring H, Sørensen HJ, Horsdal HT, MacCabe JH and Gasse C (2016*b*) Predictors of treatment resistance in patients with schizophrenia: a population-based cohort study. *The Lancet Psychiatry* **3**, 358–366.

Wimberley T, Gasse C, Meier SM, Agerbo E, MacCabe JH and Horsdal HT (2017) Polygenic risk score for schizophrenia and treatment-resistant schizophrenia. *Schizophrenia Bulletin* **43**, 1064–1069.

World Health Organization (1993) *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization.

Yoshimura B, Yada Y, So R, Takaki M and Yamada N (2017) The critical treatment window of clozapine in treatment-resistant schizophrenia: secondary analysis of an observational study. *Psychiatry Research* **250**, 65–70.

Table 18. Supplementary material published online at tiny.cc/t355ez.

On the 1st November 2017 we ran the following searches:

Pubmed:
((treatment resistant[Title/Abstract]) OR (treatment resistance[Title/Abstract]) OR (treatment refractory[Title/Abstract])) AND (schizophrenia[Title/Abstract]) AND ((longitudinal[Title/Abstract]) OR (prospective[Title/Abstract]))
Restricted to Humans and English Language
N=91 (After removing duplicates N=73)

Embase 1974 to 2017 Week 44
((treatment resistant.mp.) OR (treatment resistance.mp.) OR (treatment refractory.mp.)) AND (schizophrenia.mp.) AND ((longitudinal.mp.) OR (prospective.mp.))
Restricted to Abstracts and Humans and English Language
N=196 (After removing duplicates N=93)

Ovid MEDLINE(R) 1946 to October Week 4 2017
((treatment resistant.mp.) OR (treatment resistance.mp.) OR (treatment refractory.mp.)) AND (schizophrenia.mp.) AND ((longitudinal.mp.) OR (prospective.mp.))
Restricted to Abstracts and Humans and English Language
N=137 (After removing duplicates N=32)

PsycINFO 1806 to October Week 4 2017
((treatment resistant.mp.) OR (treatment resistance.mp.) OR (treatment refractory.mp.)) AND (schizophrenia.mp.) AND ((longitudinal.mp.) OR (prospective.mp.))
Restricted to Abstracts and Humans
N=119 (After removing duplicates N=55)

OpenGrey
((treatment resistant) OR (treatment resistance) OR (treatment refractory)) AND (schizophrenia) AND ((longitudinal) OR (prospective))
N=2 (After removing duplicates N=2)

In addition, we examined the first 20 pages of google scholar using the terms ‘predictor AND treatment resistant AND schizophrenia’ on 3rd January 2018.
https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&q=predictor+AND+treatment+resistant+AND+schizophrenia&btnG

Table 19. Modified Newcastle-Ottawa (NOS) scale for cohort studies.

Selection				**Comparability**	**Outcome**		
Representativeness of the cohort who developed the outcome (max 1)	Selection of the cohort who did not develop the outcome (max 1)	Ascertainment of exposure (max 1)	Demonstration that outcome of interest was not present at start of study (i.e. when predictors were measured) (max 1)	Comparability of cohorts on the basis of the design or analysis (max 2)	Assessment of outcome (max 1)	Was follow-up long enough for outcomes to occur (max 1)	Adequacy of follow-up of cohorts (max 1)
a) Truly representative of the average patient with psychosis in the community * b) Somewhat representative of the average patient with psychosis in the community * c) Selected group of patients with psychosis d) No description of the derivation of the cohort	a) Drawn from the same community as the exposed cohort * b) Drawn from a different source c) No description of the derivation of the non-exposed cohort	a) Secure record (e.g. clinical records) * b) Structured interview * c) Written self-report d) No description	a) Yes * b) No	a) Study controls for baseline variables correlated with the outcome * b) Study controls for any additional factors/other sensitivity analyses *	a) Independent blind assessment * b) Record linkage * c) Self report d) No description	a) Yes (1 year or longer) * b) No	a) Complete follow-up; all subjects accounted for * b) Subjects lost to follow-up unlikely to introduce bias; small number lost; description of those lost * c) Follow-up rate is large; no description of those lost d) No statement

*, Criteria which equate to 1 point

Table 20. Modified Newcastle-Ottawa (NOS) scale ratings for the studies included in this review.

Study	Representiveness of the TRS cohort	Selection of the Non-TRS cohort	Ascertainment of exposure	Demonstration that outcome of interest was not present at start of study	Comparability of cohorts on the basis of the design/analysis (max 2)	Assessment of outcome	Was follow-up long enough for outcome to occur	Adequacy of follow-up of cohorts	Score (max 9)
Chan et al. (2014)	a) Truly representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) *	a) Yes *	b) Study controls for any additional factors/other sensitivity analyses *	b) Record linkage *	a) Yes (1 year or longer) *	d) No statement	7
Demjaha et al. (2017)	a) Truly representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) * AND b) Structured interview *	a) Yes *	b) Study controls for any additional factors/other sensitivity analyses *	b) Record linkage *	a) Yes (1 year or longer) *	b) Subjects lost to follow-up unlikely to introduce bias; small number lost; description of those lost *	8
Horsdal et al. (2017a)	b) Somewhat representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) *	b) No	b) Study controls for any additional factors/other sensitivity analyses *	b) Record linkage *	a) Yes (1 year or longer) *	d) No statement	6
Horsdal, Wimberley, Köhler-Forsberg, Baandrup, and Gasse (2017b)	b) Somewhat representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) *	a) Yes *	b) Study controls for any additional factors/other sensitivity analyses *	b) Record linkage *	a) Yes (1 year or longer) *	b) Subjects lost to follow-up unlikely to introduce bias; small number lost; description of those lost *	8
Kim et al. (2017)	a) Truly representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) *	a) Yes *	b) Study controls for any additional factors/other sensitivity analyses *	b) Record linkage *	Not reported	d) No statement	6

(Lally et al., 2016a)	a) Truly representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) * AND b) Structured interview *	a) Yes *	a) Study controls for baseline variables correlated with the outcome *	b) Record linkage *	a) Yes (1 year or longer) *	b) Subjects lost to follow-up unlikely to introduce bias; small number lost; description of those lost *	8
(Meltzer et al., 1997)	c) Selected group of patients with psychosis	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) * AND b) Structured interview *	b) No	b) Study controls for any additional factors/other sensitivity analyses *	b) Record linkage *	Not reported	d) No statement	4
(Sorensen et al., 2014)	a) Truly representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) *	a) Yes *	b) Study controls for any additional factors *	b) Record linkage *	a) Yes (1 year or longer) *	b) Subjects lost to follow-up unlikely to introduce bias; small number lost; description of those lost *	8
Üçok et al. (2016)	c) Selected group of patients with psychosis	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) * AND b) Structured interview *	a) Yes *	a) Study controls for baseline variables correlated with the outcome*	b) Record linkage *	a) Yes (1 year or longer) *	b) Subjects lost to follow-up unlikely to introduce bias; small number lost; description of those lost *	7

Wimberley et al. (2016a)	a) Truly representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) *	a) Yes *	a) Study controls for baseline variables correlated with the outcome * AND b) Study controls for any additional factors/other sensitivity analyses *	b) Record linkage *	a) Yes (1 year or longer) *	b) Subjects lost to follow-up unlikely to introduce bias; small number lost; description of those lost *	9
Wimberley et al. (2016c)	a) Truly representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) *	a) Yes *	a) Study controls for baseline variables correlated with the outcome * AND b) Study controls for any additional factors/other sensitivity analyses *	b) Record linkage *	a) Yes (1 year or longer) *	b) Subjects lost to follow-up unlikely to introduce bias; small number lost; description of those lost *	9
(Wimberley et al., 2017)	b) Somewhat representative of the average patient with psychosis in the community *	a) Drawn from the same community as the exposed cohort *	a) Secure record (e.g. clinical records) *	a) Yes *	a) Study controls for baseline variables correlated with the outcome * and b) Study controls for any additional factors *	b) Record linkage *	a) Yes (1 year or longer) *	d) No statement	8

Table 21. Study Characteristics.

Study	Study Design, Year of Enrolment	N	Diagnostic Criteria	Follow-up (N)	TRS Operational Criteria	Non-TRS Operational Criteria
Chan et al. (2014)	First episode psychosis cohort recruited between Jan 1998 and Aug 2003 (Hong Kong); selected patients were prospectively followed up at interview and through electronic medical records (ratio of 1:2 TRS patients to non-TRS patients)	1400	Schizophrenia, schizoaffective disorder, acute and transient psychotic disorder, psychosis not otherwise specified	Approximately 10 years; up to May 2004 (N=474, 34%)	N=165 (35%) (1) Clozapine prescription or (2) At least moderate severity (score ≥5) on 1 or more items on the positive symptom subscale of PANSS and a score < 59 on the SOFAS at follow-up with no periods of symptomatic or functional improvement in the preceding 6 months after receiving ≥2 different types of antipsychotic medications sequentially and each for ≥4 consecutive weeks at chlorpromazine (CPZ) equivalent dose of ≥400mg/day	N=309 (65%) None
Demjaha et al. (2017)	First episode psychosis cohort recruited between Sep 1997 and Aug 2000 (UK); prospectively followed up at interview and through electronic medical records	557	ICD-10: F10-F29, F30-F33	Approximately 10 years (N=274, 49%)	N=62 (23%) (1) Persistent psychotic symptoms, defined as having a rating of at least moderate severity on one or more positive symptoms as rated by SCAN and despite recorded adherence to medication, after two sequential antipsychotic trials, each of at least 4 weeks' duration at a daily dose of 400–600 mg of chlorpromazine equivalents	N=212 (77%) (1) A period of at least 6 months' duration in which no symptoms or only symptoms of mild severity, not interfering with daily functioning, were experienced

Horsdal et al. (2017a)	First diagnosis schizophrenia cohort curated from electronic medical records of patients born in Denmark after 1954, who had a baseline measure of C-reactive protein (CRP), and whose first diagnosis of schizophrenia occurred between Feb 2000 and Nov 2012; prospectively followed up through electronic records	390	ICD-8: 295.x9 excluding 295.79 or ICD-10: F20	2 years; the follow-up period ran from first diagnosis of schizophrenia until incidence of TRS, emigration, death, or 2 years after first schizophrenia diagnosis, whichever came first (N=390, 100%)	N=52 (13%) (1) First clozapine prescription redemption or (2) Psychiatric hospital admission due to schizophrenia during antipsychotic treatment within the 18 months after having received two prior antipsychotic monotherapy trials of adequate duration (6 weeks); hospitalizations from the year prior to first diagnosis and until the study endpoint were included	N=338 (87%) None
Horsdal et al. (2017b)	First diagnosis schizophrenia cohort curated from electronic medical records of patients born in Denmark after 1954, who had a baseline measure of the GAF-F, and whose first diagnosis of schizophrenia occurred between Jan 2004 and Dec 2010; prospectively followed up through electronic records	3252	ICD-10: F20	2 years; the follow-up period ran from first diagnosis of schizophrenia until incidence of TRS, emigration, death, or 2 years after first schizophrenia diagnosis, whichever came first (N=3252, 100%)	N=359 (11%) (1) First clozapine prescription redemption or (2) Psychiatric hospital admission due to schizophrenia during antipsychotic treatment within the 18 months after having received two prior antipsychotic monotherapy trials of adequate duration (6 weeks); hospitalizations from the year prior to first diagnosis and until the study endpoint were included	N=2893 (89%) None
Kim et al. (2017)	New antipsychotic-users cohort curated from electronic medical records of patients between Jan 2008 and Dec 2014 (South Korea); prospectively followed up through electronic records	114,749	ICD-10: F20	Not reported; for some patients follow-up exceeded 6 years	N not reported (1) First prescription of clozapine	N not reported None

(Lally et al., 2016a)	First episode psychosis cohort recruited between Dec 2005 and Oct 2010 (UK); prospectively followed up through electronic medical records	283	ICD-10: F20.0, F25.0, F28.0, F29.0	5 years (N=240, 85%)	N=81 (34%) (1) Treated with clozapine at any point during the follow-up period or (2) Little or no symptomatic improvement to two consecutive treatments with antipsychotic medications of adequate dose (400 mg chlorpromazine equivalence) and duration (at least 6 weeks) excluding those intolerant to antipsychotic medications or those who self-discontinued medication	N=159 (66%) None
(Meltzer et al., 1997)	First hospitalisation for schizophrenia cohort (USA), recruitment period not reported, prospectively followed up through electronic medical records	322	DSM-III-R; schizophrenia, schizoaffective disorder	Not reported; for some patients follow-up exceeded 4 years	N= 196 (61%) (1) Persistent moderate to severe delusions, hallucinations, or thought disorder, despite at least three trials of typical neuroleptic drugs for at least 6 weeks at adequate doses or (2) Pervasive negative symptoms, such as withdrawal, anhedonia, poverty of thought content, a deficit in volition, and lack of energy, despite at least three trials of typical neuroleptic drugs for at least 6 weeks at adequate doses	N= 126 (39%) (1) Those who had at most mild positive and negative symptoms during the most recent course of neuroleptic treatment
(Sorensen et al., 2014)	First diagnosis schizophrenia cohort curated from electronic medical records of patients born in Denmark between 1950 and 1970 whose first diagnosis of schizophrenia occurred between 1975 and 1990; prospectively followed up through electronic records	5968	ICD-8: 295	5-34 years; follow-up to first clozapine prescription, date of death, or 31st December 2009, whichever came first (N=5328, 89%)	N=1223 (23%) (1) First prescription of clozapine between 1995 to 2009	N=4105 (77%) None

Üçok et al. (2016)	First episode psychosis cohort recruited between 1996 and 2016 (Turkey); prospectively followed up at interview	187	DSM-IV: schizophrenia	Not reported; minimum 6 months for TRS patients and 2 years for non-TRS patients (N=105, 56%)	N=28 (27%) (1) Clozapine prescription	N=777 (3%) None
Wimberley et al. (2016a)	First diagnosis schizophrenia cohort curated from electronic medical records of patients born in Denmark after 1955 and whose first diagnosis of schizophrenia occurred between Jan 1996 and Jul 2013; prospectively followed up through electronic records	13,349	ICD-10: F20	Median of 7 (IQR: 3-12) years; individuals were followed from their first diagnosis of schizophrenia until they met criteria for TRS, emigrated from Denmark, died, or until 1st July 2013, whichever came first (N=13,349, 100%)	N=2313 (17%) (1) Clozapine prescription or (2) Psychiatric hospital admission due to schizophrenia, with evidence of treatment adherence, after having received two prior antipsychotic monotherapy trials of adequate duration (6 weeks), counted from one year prior to the first recorded schizophrenia diagnosis	N=11,036 (83%) None
Wimberley et al. (2016c)	First diagnosis schizophrenia cohort curated from electronic medical records of patients born in Denmark after 1955 and whose first diagnosis of schizophrenia occurred between Jan 1996 and Dec 2006; prospectively followed up through electronic records	9332	ICD-8: 295.x9 (excluding 295.79) or ICD-10: F20	Median 9 (IQR: 6-12) years; individuals were followed from their first diagnosis of schizophrenia until emigration, death, or 31st Dec 2010, whichever came first (N=8044, 86%)	N=1703 (21%) (1) Clozapine prescription or (2) Psychiatric hospital admission due to schizophrenia during antipsychotic treatment within the 18 months after having received two prior antipsychotic monotherapy trials of adequate duration (6 weeks), counted from one year prior to the first recorded schizophrenia diagnosis	N=6341 (79%) None

(Wimberley et al., 2017)	First diagnosis schizophrenia cohort curated from electronic medical records of patients born in Denmark after 1981, who had a DNA sample available, and whose first diagnosis of schizophrenia occurred between 1999 and 2007; prospectively followed up through electronic records	862	ICD-10: F20	Median 5 (IQR: 4-7) years; individuals were followed from their first diagnosis of schizophrenia until incidence of TRS, emigration, death, or 31st Dec 2010, whichever came first (N=862, 100%)	N=181 (21%) (1) Clozapine prescription or (2) Psychiatric hospital admission due to schizophrenia during antipsychotic treatment within the 18 months after having received two prior antipsychotic monotherapy trials of adequate duration (6 weeks), counted from one year prior to the first recorded schizophrenia diagnosis	N=681 (79%) None

Abbreviations: DNA, deoxyribonucleic acid; DSM, Diagnostic Statistical Manual; GAF-F, Global Assessment of Functioning – functioning scale; ICD, International Classification of Diseases; IQR, interquartile range; PANSS, Positive and Negative Schizophrenia Symptom scale; SCAN, Schedules for Clinical Assessment in Neuropsychiatry, SCAN; SOFAS, Social and Occupational Functioning Assessment Scale; TRS, treatment resistant schizophrenia.

Table 22. Unadjusted and adjusted OR/HR, when reported, for the studies included in this review.

Predictors	Study	Unadjusted		Adjusted	
		OR/HR	95%CI	OR/HR	95%CI
Patient Characteristics					
Born Autumn (September–November) vs. spring (March–May)	Sorensen et al. (2014) †	NR	NR	1.24	1.06–1.46
Born December–March vs April–November	Wimberley *et al.* (2016b) †	NR	NR	1.01	0.91–1.11
Born Summer (June–August) vs. spring (March–May)	Sorensen et al. (2014)†	NR	NR	1.15	0.98–1.35
Born Winter (December–February) vs. spring (March–May)	Sorensen et al. (2014)†	NR	NR	1.12	0.95–1.31
CRP >3 mg/L vs. CRP 0–3 mg/L	Horsdal et al. (2017a)†	1.14	0.67–1.96	0.99	0.56–1.73
Early parental loss (<18 years)	Wimberley *et al.* (2016b)†	NR	NR	0.97	0.81–1.16
Education: primary vs higher education level	Wimberley *et al.* (2016b)†	NR	NR	0.88	0.79–0.98
Education: years of education	Chan *et al.* (2014)†	NR	NR	0.98	0.92–1.04
Employment: long–term disability benefit vs. in work	Wimberley *et al.* (2016b)†	NR	NR	1.14	0.95–1.36
Employment: outside working force vs. in work	Wimberley *et al.* (2016b)†	NR	NR	1.01	0.90–1.13
Employment: unemployed vs. in work	Wimberley *et al.* (2016b)†	NR	NR	0.95	0.79–1.15
Family history of schizophrenia vs no family history	Wimberley *et al.* (2016b)†	NR	NR	1.00	0.83–1.21
Female vs male	Wimberley *et al.* (2016b)†	NR	NR	1.07	0.96–1.19
Female vs. male	Demjaha *et al.* (2017)	0.48	0.25–0.86	NR	NR
Level of urbanicity (diagnosis): provincial areas vs. capital area	Wimberley et al. (2016a)†	NR	NR	1.40	1.26–1.56
Level of urbanicity (diagnosis): rural areas vs. capital area	Wimberley et al. (2016a)†	NR	NR	1.56	1.39–1.76
Living alone vs in a couple	Wimberley *et al.* (2016b)†	NR	NR	1.00	0.91–1.11
Non–white vs. white	Demjaha *et al.* (2017)	1.26	0.72–2.24	NR	NR
PAS adult (19+ years) subscale score	Chan *et al.* (2014)†	NR	NR	3.22	1.43–7.23
Paternal age	Wimberley *et al.* (2016b) †	NR	NR	1.00	0.99–1.01
Polygenic risk score	Wimberley et al. (2017) †	1.09	0.92–1.30	1.13	0.95–1.35
Urbanicity: provincial area (>10 000 inhabitants) vs. capital area (capital and suburb to the capital)	Wimberley *et al.* (2016b) †	NR	NR	1.38	1.23–1.56
Urbanicity: rural area vs. capital area (capital and suburb to the capital)	Wimberley *et al.* (2016b) †	NR	NR	1.44	1.25–1.65
Substance misuse history	Chan *et al.* (2014) †	NR	NR	1.34	0.67–2.69
Violent offence	Wimberley *et al.* (2016b) †	NR	NR	1.04	0.89–1.23
Disease Characteristics					
>30 bed–days in psychiatric hospital in year before first schizophrenia diagnosis vs. 0 bed–days	Wimberley *et al.* (2016b) †	NR	NR	1.54	1.35–1.75
1–30 bed–days in psychiatric hospital in year before first schizophrenia diagnosis vs. 0 bed–days	Wimberley *et al.* (2016b) †	NR	NR	1.11	0.96–1.27
Age at first schizophrenia diagnosis	Wimberley *et al.* (2016b) †	NR	NR	0.96	0.95–0.97
Age of onset	Chan *et al.* (2014) †	NR	NR	0.88	0.83–0.94
Age of onset	Demjaha *et al.* (2017)	0.93	0.89–0.97	0.97	NR
Age of onset	Üçok et al. (2016)	NR	NR	1.018	NR

Age of onset <20 years	Lally *et al.* (2016)	NR	NR	2.49	1.25–4.94
Age of onset >31 years	Lally *et al.* (2016)	NR	NR	0.55	0.27–1.12
Age of onset 21–25 years	Lally *et al.* (2016)	NR	NR	0.70	0.37–1.32
Age of onset 26–30 years	Lally *et al.* (2016)	NR	NR	1.14	0.57–2.28
Diagnosis: depressive psychosis vs. schizophrenia	Demjaha *et al.* (2017)	0.15	0.02–0.50	0.52	NR
Diagnosis: manic psychosis vs. schizophrenia	Demjaha *et al.* (2017)	0.12	0.02–0.42	0.41	NR
Diagnosis: Paranoid subtype	Wimberley *et al.* (2016b) †	NR	NR	1.24	1.13–1.37
Diagnosis: previous comorbid comorbid diagnosis of suicide attempt	Wimberley *et al.* (2016b) †	NR	NR	1.21	1.07–1.39
Diagnosis: previous comorbid comorbid diagnosis of other schizophrenia spectrum disorders	Wimberley *et al.* (2016b) †	NR	NR	1.05	0.94–1.16
Diagnosis: previous comorbid comorbid diagnosis of personality disorder	Wimberley *et al.* (2016b) †	NR	NR	1.24	1.11–1.39
Diagnosis: previous comorbid comorbid diagnosis of schizoaffective disorder	Wimberley *et al.* (2016b) †	NR	NR	1.18	0.95–1.45
Diagnosis: previous comorbid comorbid diagnosis of substance abuse	Wimberley *et al.* (2016b) †	NR	NR	0.99	0.88–1.11
Diagnosis: previous comorbid diagnosis of depression	Wimberley *et al.* (2016b) †	NR	NR	1.11	0.97–1.26
Duration of first episode	Chan *et al.* (2014) †	NR	NR	1.003	1.001–1.004
Duration of untreated psychosis (days)	Chan *et al.* (2014) †	NR	NR	0.90	0.69–1.18
Duration of untreated psychosis (weeks)	Demjaha *et al.* (2017)	1.004	1.001–1.006	1.0013	NR
Duration of untreated psychosis (days)	Üçok et al. (2016)	NR	NR	1.03	NR
GAF-F	Lally *et al.* (2016)	NR	NR	0.98	0.96–1.01
GAF-F score 1–30 vs. 31–100	Horsdal et al. (2017b) †	1.45	1.08–1.94	1.38	1.03–1.86
GAF-S	Lally *et al.* (2016)	NR	NR	0.97	0.94–1.00
Inpatient at first schizophrenia diagnosis	Wimberley *et al.* (2016b) †	NR	NR	2.07	1.87–2.29
Mode of onset: insidious vs. acute	Demjaha *et al.* (2017)	0.13	1.44–0.00	1.28	NR
Negative symptoms (derived using FA from the SCAN)	Demjaha *et al.* (2017)	1.24	1.08–1.42	1.09	NR
PANSS Conceptual disorganization	Lally *et al.* (2016)	NR	NR	1.13	0.83–1.56
PANSS Lack of judgement and insight	Lally *et al.* (2016)	NR	NR	1.14	0.89–1.47
PANSS Negative	Lally *et al.* (2016)	NR	NR	1.03	0.96–1.11
PANSS Positive	Lally *et al.* (2016)	NR	NR	1.01	0.94–1.07
PANSS Total	Lally *et al.* (2016)	NR	NR	1.01	0.98–1.04
Psychotropic medication redeemed in year before first schizophrenia diagnosis: antidepressants	Wimberley *et al.* (2016b) †	NR	NR	1.15	1.03–1.29
Psychotropic medication redeemed in year before first schizophrenia diagnosis: antipsychotics	Wimberley *et al.* (2016b) †	NR	NR	1.51	1.35–1.69
Psychotropic medication redeemed in year before first schizophrenia diagnosis: benzodiazepines	Wimberley *et al.* (2016b) †	NR	NR	1.22	1.10–1.37
During Treatment					
Antipsychotic polypharmacy (prior to clozapine for clozapine users)	Üçok et al. (2016)	NR	NR	0.15	NR
Having a first relapse despite being adherent to non–clozapine antipsychotic treatment	Üçok et al. (2016)	NR	NR	3.93	NR
Number of relapses in the first three years	Chan *et al.* (2014) †	NR	NR	1.45	1.18–1.78
Relapse in first 6 months (prior to clozapine use for clozapine users)	Üçok et al. (2016)	NR	NR	1.03	NR

NB: Kim et al. (2017) and Meltzer et al. (1997) did not report OR/HR and therefore not included in this table.
Abbreviations: †, Hazard Ratios; CRP, C-reactive protein; FA, factor analysis; GAF-F, Global Assessment of Functioning – functioning scale; GAF-S, Global Assessment of Functioning – symptoms scale; PANSS, Positive and Negative Schizophrenia Symptom scale; PAS, Premorbid Assessment Scale; SCAN, Schedules for Clinical Assessment in Neuropsychiatry.

Table 23. PRISMA Checklist.

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4-5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	No
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4-5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5/Supplementary material
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6/Supplementary material
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	NA

Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	14-17
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	6/Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	6-11/Table 1/Table 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Supplementary material
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	6-11/Table 1/Table 2/ Supplementary material
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	NA
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-16
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14-16
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	11-16
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	1

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

14.3. Appendix C

Table 24. Apparent coefficients for the LASSO logistic regression.

	Original Model		Using clozapine vs. no clozapine as the outcome		Restricted to those with a diagnosis of schizophrenia at follow-up		Oversampling	
Predictor	**Coefficient**	**Direction of Effect**	**Coefficient**	**Direction of Effect**	**Coefficient**	**Direction of Effect**	**Coefficient**	**Direction of Effect**
Intercept	0.32		-1.26		-1.80		0.83	
Age of onset (years)	-0.02	-	-0.05	-	-0.02	-	-0.03	-
DUP (days)			-0.0001	-			0.0001	+
Female (vs. male)	-0.10	-	0.01	+	-0.17	-	-0.18	-
Body Mass Index	0.002	+	0.02	+	0.002	+	0.03	+
Being in a relationship (vs. single)	-0.15	-	-0.30	-	-0.37	-	-0.36	-
Living alone (vs. with others)								
Living with non-parents (vs. with parents)	0.06	+	0.40	+			0.14	+
Years in Education	-0.04	-	-0.01	-			-0.01	-
Cannabis (vs. no cannabis)							-0.13	-
Non-smoker (vs. smoker)							0.03	+
Non-drinker (vs. drinker)	-0.33	-	-0.22	-			-0.10	-
PANSS Positive							0.001	+
PANSS Negative			0.04	+	0.03	+	0.02	+
PANSS General Psychopathology								
SAPS					0.04	+	0.01	+
SANS					0.003	+		
GAF	-0.02	-	-0.01	-	-0.004	-	-0.02	-
No qualifications (vs. all other qualifications)								
Basic education qualification (vs. all other qualifications)							-0.11	-
Further education qualification (vs. all other qualifications)	0.16	+	-0.08	-			-0.19	-
Higher education qualification (vs. all other qualifications)			-0.03	-			-0.57	-
Non-European (vs. European)								
Black (vs. non-black)			0.39	+			-0.08	-
Asian/Other (vs. non-Asian/Other)			0.17	+			-0.13	-

Table 25. Apparent performance measures for the four LASSO logistic regressions: the original model, when using clozapine vs. no clozapine as the outcome, restricted to those with a diagnosis of schizophrenia at follow-up, and when oversampling TR cases.

	Original Model		**Using clozapine vs. no clozapine as the outcome**		**Restricted to those with a diagnosis of schizophrenia at follow-up**		**Oversampling**	
Performance Measure	**50% Threshold**	**19.1% Threshold**	**50% Threshold**	**10.8% Threshold**	**50% Threshold**	**21.8% Threshold**	**50% Threshold**	**48.5% Threshold**
Alpha	0.47		0.53		1.10		-0.001	
Beta	1.32		1.28		1.79		1.09	
AUC	0.65		0.67		0.66		0.67	
Accuracy	0.83	0.64	0.88	0.50	0.81	0.69	0.62	0.63
Sensitivity (%)	0	58.87	0	80.44	0	51.82	63.30	70.37
Specificity (%)	100	64.72	100	45.81	100	72.77	59.90	55.49
Positive Predictive Value (%)	NA	26.01	NA	16.56	NA	30.81	61.22	61.25
Negative Predictive Value (%)	82.60	88.19	88.21	94.60	81.03	86.58	62.01	65.19

Abbreviations: AUC, area under the receiver operating characteristic curve; NA, not applicable.
NB 1: When using clozapine vs. no clozapine as the outcome, the sample size was N=1908, of which N=1683 (88.21%) were responders and N=225 (11.79%) were TR, and 14 variables were selected. Lambda = 0.003854.
NB 2: When using only those with a diagnosis of schizophrenia at follow-up the sample size was N=580, of which N=470 (81.03%) were responders and N=110 (18.97%) were TR, and 8 variables were selected. Lambda = 0.017074.
NB 3: When oversampling TR, the sample size was N=3172, of which N= 1586 (50%) were responders and N=1586 (50%) were TR, and 20 variables were selected. Lambda = 0.003511.
NB 4: The higher the alpha, the more coefficients are 0 in the model. When alpha = 0, the model is the same as logistic regression. When beta > 1 the model is underfitting the data, when beta < 1 the model is overfitting the data.

Table 26. Performance measures for the original LASSO logistic regressions, stratified by cohort, after repeated cross-fold validation.

Cohort	**Performance Measures**		
	Alpha	**Beta**	**AUC**
AESOP London	1.98	1.88	0.64
Belfast	-5.50	1.77	0.66
Bologna	-9.53	6.85	0.70
GAP London	0.51	1.11	0.60
Istanbul	-0.38	0.20	0.65
Lausanne	1.67	2.75	0.64
Paris	226.92	180.17	0.76
Prague	14.63	10.05	0.67
Santander	1.13	1.86	0.63
UCL London	1.22	3.15	0.67

Table 27. Recalibrated coefficients for the LASSO logistic regression: the original model, when using clozapine vs. no clozapine as the outcome, restricted to those with a diagnosis of schizophrenia at follow-up, and when oversampling TR cases.

	Original Model		**Using clozapine vs. no clozapine as the outcome**		**Restricted to those with a diagnosis of schizophrenia at follow-up**		**Oversampling**	
Predictor	**Coefficient**	**Direction of Effect**	**Coefficient**	**Direction of Effect**	**Coefficient**	**Direction of Effect**	**Coefficient**	**Direction of Effect**
Intercept	0.26		-1.33		-1.76		0.75	
Age of onset (years)	-0.02	-	-0.04	-	-0.02	-	-0.03	-
DUP (days)			-0.00008	-			0.0001	+
Female (vs. male)	-0.09	-	0.01	+	-0.17	-	-0.18	-
Body Mass Index	0.002	+	0.02	+	0.002	+	0.03	+
Being in a relationship (vs. single)	-0.14	-	-0.27	-	-0.36	-	-0.35	-
Living alone (vs. with others)								
Living with non-parents (vs. with parents)	0.06	+	0.36	+			0.14	+
Years in Education	-0.03	-	-0.01	-			-0.01	-
Cannabis (vs. no cannabis)							-0.12	-
Non-smoker (vs. smoker)							0.03	+
Non-drinker (vs. drinker)	-0.31	-	-0.19	-			-0.09	-
PANSS Positive							0.001	+
PANSS Negative			0.04	+	0.02	+	0.02	+
PANSS General Psychopathology								
SAPS					0.04	+	0.01	+
SANS					0.003	+		
GAF	-0.02	-	-0.01	-	-0.004	-	-0.02	-
No qualifications (vs. all other qualifications)								
Basic education qualification (vs. all other qualifications)							-0.11	-
Further education qualification (vs. all other qualifications)	0.15	+	-0.07	-			-0.19	-
Higher education qualification (vs. all other qualifications)			-0.03	-			-0.56	-
Non-European (vs. European)			0.35	+				
Black (vs. non-black)			0.15	+			-0.08	-
Asian/Other (vs. non-Asian/Other)							-0.13	-

Table 28. Corrected performance measures for the four LASSO logistic regressions: the original model, when using clozapine vs. no clozapine as the outcome, restricted to those with a diagnosis of schizophrenia at follow-up, and when oversampling TR cases.

	Original Model		**Using clozapine vs. no clozapine as the outcome**		**Restricted to those with a diagnosis of schizophrenia at follow-up**		**Oversampling**	
Performance Measure	**50% Threshold**	**19.1% Threshold**	**50% Threshold**	**10.8% Threshold**	**50% Threshold**	**21.8% Threshold**	**50% Threshold**	**48.5% Threshold**
Alpha	-0.05		-0.21		-0.04		-0.06	
Beta	0.96		0.89		0.95		0.98	
AUC	0.60		0.62		0.56		0.64	
Accuracy	0.76	0.57	0.87	0.48	0.72	0.60	0.61	0.62
Sensitivity (%)	00.01	58.88	00.00	80.45	1.76	52.57	66.16	73.22
Specificity (%)	99.98	64.71	100	45.81	99.24	72.01	53.25	48.84
Positive Predictive Value (%)	NA	24.13	NA	16.40	NA	25.58	58.87	59.60
Negative Predictive Value (%)	82.60	88.47	88.21	94.34	81.04	84.92	61.40	66.11

Abbreviations: AUC, area under the receiver operating characteristic curve; NA, not applicable.
NB 1: When using clozapine vs. no clozapine as the outcome, the sample size was N=1908, of which N=1683 (88.21%) were responders and N=225 (11.79%) were TR, and 14 variables were selected.
NB 2: When using only those with a diagnosis of schizophrenia at follow-up the sample size was N=580, of which N=470 (81.03%) were responders and N=110 (18.97%) were TR, and 8 variables were selected.
NB 3: When oversampling TR, the sample size was N=3172, of which N= 1586 (50%) were responders and N=1586 (50%) were TR, and 20 variables were selected.
NB 4: The higher the alpha, the more coefficients are 0 in the model. When alpha = 0, the model is the same as logistic regression. When beta > 1 the model is underfitting the data, when beta < 1 the model is overfitting the data.

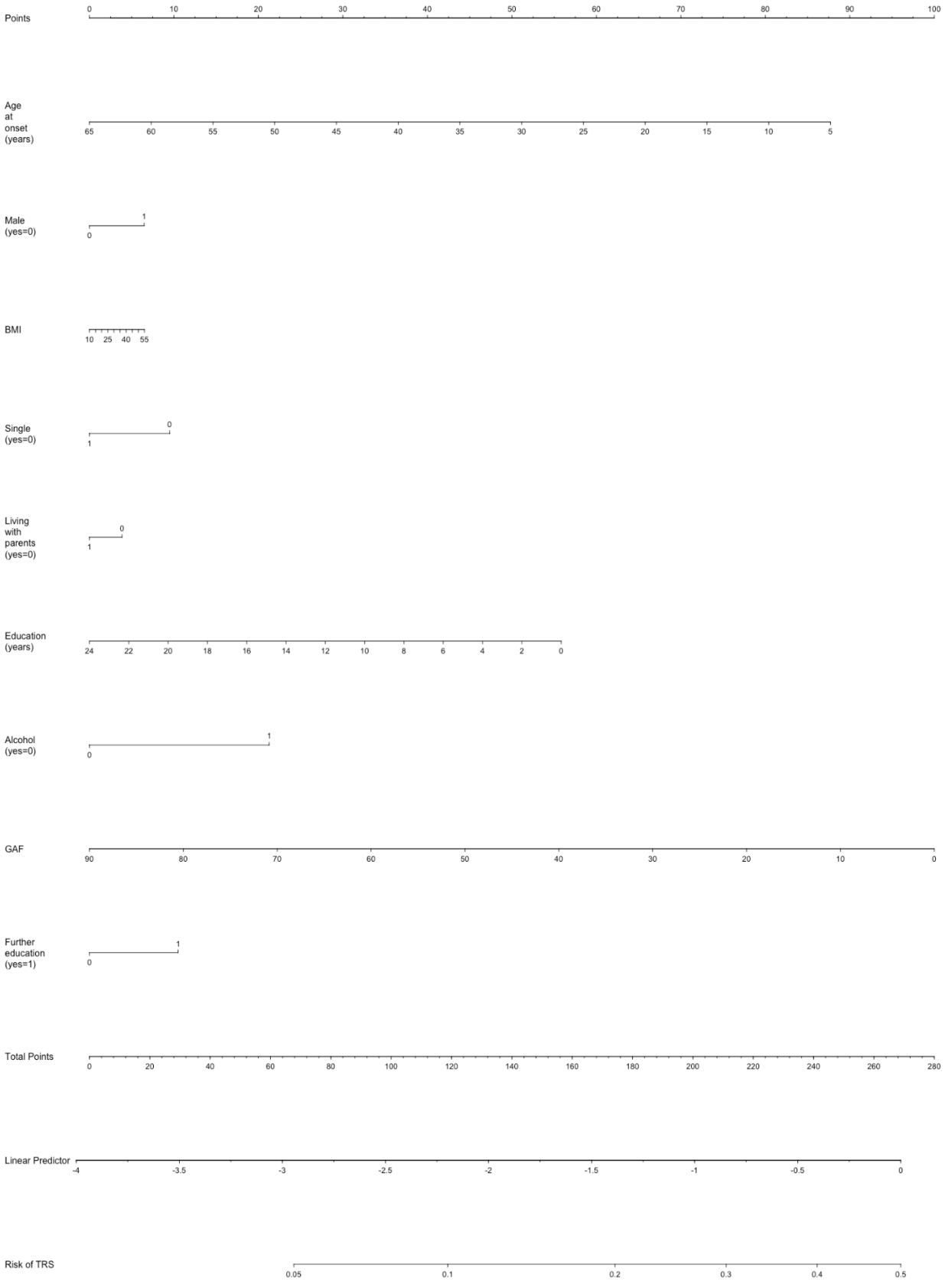
Equation 1. The equation to predict TR using 1SE LASSO logistic regression model.

$\text{logit}(\hat{p})$ provides the probability of being TR for each new observation. If the probability is over 19.1%, then the model would classify the observation as TR. If a patient's probability of being TR is 90%, a clinician could be very certain that they will not respond to conventional antipsychotic medication. If a patient's probability of being TR is 20%, the model would classify them as TR, but a clinician may decide that treatment as usual is more appropriate than any alternative intervention.

$$\text{logit}(\hat{p}) = \log\left(\frac{\hat{p}}{1-\hat{p}}\right)$$
$$= 0.26 - 0.021397734930248 \times \text{age of onset (years)}$$
$$- 0.0948587262403061 \times \text{gender (female)}$$
$$+ 0.00211790080461775 \times \text{BMI}$$
$$- 0.138955209285684 \times \text{being in a relationship (yes)}$$
$$+ 0.0564345279638353 \times \text{living with nonparents (yes)}$$
$$- 0.0340555064203194 \times \text{education (years)}$$
$$- 0.311338300241959 \times \text{alcohol (no)}$$
$$- 0.0162638785427521 \times \text{GAF}$$
$$+ 0.153672741276086$$
$$\times \text{further education qualification (yes)}$$

Figure 10. Nomogram plot to predict TR in new cases.

To predict TR, one can take the example of a 20-year-old male patient, with a BMI of 20, who is single and lives with their parents, spent 12 years in education, drinks, scores 20 on the GAF scale, and who has A-levels qualifications. They would have a total score of 66 + 6 + 1 + 9 + 4 + 28 + 0 + 78 + 10 = 202. This corresponds to a ~28% risk of developing TR.


Points
0 10 20 30 40 50 60 70 80 90 100
Age at onset (years)
65 60 55 50 45 40 35 30 25 20 15 10 5
Male (yes=0)
1
0
BMI
10 25 40 55
Single (yes=0)
0
1
Living with parents (yes=0)
0
1
Education (years)
24 22 20 18 16 14 12 10 8 6 4 2 0
Alcohol (yes=0)
1
0
GAF
90 80 70 60 50 40 30 20 10 0
Further education (yes=1)
1
0
Total Points
0 20 40 60 80 100 120 140 160 180 200 220 240 260 280
Linear Predictor
-4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0
Risk of TRS
0.05 0.1 0.2 0.3 0.4 0.5

14.4. Appendix D

A genetic interaction analysis for treatment resistance – by Dr Antonio Pardiñas

Abstract

Importance

Treatment-resistant schizophrenia (TRS) is a chronic and highly disabling condition affecting 30% of schizophrenia patients. While there is an ongoing and frequent debate about the nature of TRS, and its relationship with the biological basis of schizophrenia, adequately powered genetic studies are scarce due to the difficulty in collecting data from well-characterised TRS patients.

Objective

To examine the genetic architecture of TRS through the reassessment of genetic data from schizophrenia studies and its validation in carefully ascertained clinical samples.

Design, Setting, and Participants

We combined data from the CLOZUK and Psychiatric Genomics Consortium (PGC) schizophrenia studies, for a total of 83,841 participants (10,501 TRS; 21,264 non-TRS; 52,076 healthy controls). We used a GWAS meta-analysis approach to estimate effect size differences between TRS and non-TRS samples. We validated this procedure using polygenic risk score (PRS) profiling of two independent UK-based schizophrenia cohorts: CardiffCOGS (a prevalence sample with 341 TRS and 449 non-TRS) and STRATA-G (an incidence sample with 148 TRS and 911 non-TRS).

Main Outcomes and Measures

TRS emerged as a polygenic trait with detectable heritability ($h^2_{SNP} = 0.014 \pm 0.005$), and the corresponding TR-PRS was significantly associated with TRS status in external samples.

Results

The TRS GWAS showed one genome-wide significant signal at the known NT5C2 schizophrenia locus ($P = 5.26 \times 10^{-9}$). Several traits, including adult intelligence and smoking status, were found to be genetically correlated to TRS, with a co-localisation analysis highlighting likely pleiotropic loci. PRS analysis in CardiffCOGS showed a positive association between the TR-PRS and several TRS definitions ($r^2 = 2.70\%$; $P = 1.42 \times 10^{-4}$), which was replicated in STRATA-G ($r^2 = 1.11\%$; $P = 0.046$). Notably, summary statistics generated from the CLOZUK component of the TRS GWAS, composed of samples collected from clozapine clinics, were also predictive of TRS status in both samples. The PGC component, mostly non-TRS, was not. Other genetically correlated traits, such as educational attainment or intelligence, achieved a marginal (negative) association with TRS.

Conclusions and Relevance

Common genetic variants associated to TRS exist, and they are also related to traits (intelligence, smoking) that have been frequently discussed as of functional relevance for TRS patients. The effect of these variants might have been masked through the amalgamation of large GWAS samples. Thus, this study supports a focus on detailed phenotypes for the future genetic research of schizophrenia subtypes, and confirms the validity of this approach for TRS studies.

Introduction

Nearly 30% of schizophrenia patients do not respond adequately to conventional antipsychotic medication (Lally, Gaughran, Timms, & Curran, 2016b), and thus are considered to experience 'treatment-resistant schizophrenia' (TRS), a severe and largely chronic condition. These patients are among the most disabled of all those who experience psychiatric disorders, with frequent re-hospitalisation, poor symptomatic and functional outcomes, and decreased life expectancy due to major physical health problems and increased suicide rates (Conley & Kelly, 2001). The only indicated and licensed drug for TRS patients is clozapine, which is effective in around 60% of cases (Kumra et al., 2008) and improves most morbidity and mortality indicators (Tiihonen, Mittendorfer-Rutz, Majak, & et al., 2017a). While the mechanism of action of clozapine is still not fully understood (Lally & MacCabe, 2015), it has been proposed that its efficacy in the context of TRS might

be related to the biological basis of treatment resistance, which might implicate different neuronal systems to the majority of non-TRS cases (Krivoy et al., 2018).

While the biology of schizophrenia is increasingly becoming better understood thanks to genetic studies led by international consortia (Sullivan et al., 2017), there is considerable heterogeneity in findings related to TRS. For example, there is conflicting evidence regarding the existence of a common-variant signal in TRS patients, related to the schizophrenia polygenic score (Frank et al., 2015a; Wimberley et al., 2017), and a recent systematic review did not find any individual genes robustly associated to TRS (Gillespie et al., 2017). This is a potential reflection of the difficulties of recruiting large numbers of treatment-resistant patients for research purposes (Taylor, 2017), which has precluded the execution of well-powered genome-wide association studies (GWAS).

In this study, we aim to characterise the common variant genetic architecture of TRS by exploiting the large amounts of data generated by recent GWAS of schizophrenia (Pardiñas et al., 2018; Ripke et al., 2014), in which individuals that can be formally defined as TRS were present. Recent analyses have employed a variety of statistical techniques to assess the effects of genetic variants across health-related traits (Cleynen et al., 2016; Justice et al., 2017; Ruderfer et al., 2018), taking advantage of case-case study designs. Since a large-scale analysis of TRS against non-TRS individuals has not yet been carried out, we have adapted a meta-analytic procedure to assess the differences between GWAS in which TRS and non-TRS individuals have been compared to healthy controls. Our main hypothesis was that these differences reflect the underlying genetic architecture of TRS and can be analysed in ways analogous to other genome-wide summary statistics in order to reveal commonalities with other complex traits. This can recapitulate biological and epidemiological characteristics of this disorder, which can in turn inform about its distinctions with other forms of schizophrenia (Gillespie et al., 2017). We also propose that genetic variants associated with TRS in GWAS samples could be validated in clinical cohorts using the polygenic risk score approach (PRS) (Wray et al., 2014). In light of the evidence showing that the majority of patients who develop treatment resistance might do so close to the illness onset (Lally et al., 2016a), PRS calculated in cohorts in which TRS status

has been ascertained in different ways should be similarly powered and informative of whether a common genetic basis might exist across the heterogeneous spectrum of TRS.

Methods

GWAS Samples

We used CLOZUK1 and CLOZUK2 as samples of TRS individuals, which have been described in previous studies (Hamshere et al., 2013; Pardiñas et al., 2018). All the individuals with schizophrenia collected in these samples were prescribed clozapine in the UK after a failure to respond to two trials of common antipsychotics, following NICE guidelines for TRS (Mortimer, Singh, Shepherd, & Puthiryackal, 2010). The use of a history of taking clozapine as equivalent to a research diagnosis of TRS has been validated in these samples (Pardiñas et al., 2018), as well as in independent studies (Ekholm et al., 2005; Jakobsen et al., 2005). Control individuals, not screened for psychiatric disorders, were collected from public databases or through collaboration with population sequencing projects in the UK. In order to ameliorate effects of population stratification, the analysed individuals were those selected by Pardiñas et al. (2018), which amount to 10,501 TRS cases and 24,542 healthy controls. These were inferred to be predominantly of UK genetic ancestry after principal component analysis (PCA) and ADMIXTURE estimation, and contain no detectable population outliers.

As samples of non-TRS individuals, we used 34 studies included in the meta-analysis by the Schizophrenia Working Group of the Psychiatric Genomic Consortium (Ripke et al., 2014). From these studies, we removed all the TRS individuals we could identify based on existing clinical records. Control individuals in this analysis were a mixture of publicly available unscreened samples and clinically-ascertained individuals, and extensive analyses to discard population outliers and assess stratification were carried out Ripke et al. (2014). The resulting sample contained 21,264 schizophrenia cases and 27,534 healthy controls. Since TRS is an under-reported condition, it is possible that some treatment-resistant individuals might remain in these datasets, though they will be a minority.

Polygenic validation samples

For additional downstream analyses, we used a sample of individuals with schizophrenia that included a mixture of treatment-resistant and treatment-responsive cases, CardiffCOGS (N=790; 341 TRS and 449 non-TRS). This is a cross-sectional cohort with clinical ratings based on contemporaneous records, which was also described in a previous study (Rees et al., 2014). We also used a novel multi-ethnic sample of people with first episode psychosis, STRATA-G (N=1,059; 148 TRS and 911 non-TRS), which has been followed up for a number of years to ascertain treatment resistance. This sample, its follow-up and rating system is extensively described in [Chapter 7].

Genotyping, imputation and quality control

Imputed genotypes from the GWAS samples were extracted from their respective studies: treatment-resistant genotypes from CLOZUK1 and CLOZUK2 were retrieved from Pardiñas et al. (2018) while treatment-responsive genotypes were retrieved from Ripke et al. (2014). While all samples were imputed using the SHAPEIT/IMPUTE2 (Delaneau, Zagury, & Marchini, 2013; Howie, Fuchsberger, Stephens, Marchini, & Abecasis, 2012) algorithms, different versions of the 1000 Genomes reference panel were used in the two studies (Phase 3 and Phase 1 respectively). However, use of either reference has been found to result in similar imputation accuracies for SNPs with common (>5%) allele frequencies (1000 Genomes Project Consortium, 2015). Thus, in order to eliminate any bias due to differential imputation procedures, we restricted to SNPs with a minor allele frequency (MAF) of 5% or higher in all our analyses. Additionally, the same post-imputation filters were applied to all remaining SNPs (INFO > 0.6; Hardy Weinberg Equilibrium P-value > 10^{-6}).

Genotypes from the CardiffCOGS samples were collected and curated as described on Pardiñas et al. (2018), while the collection and curation of STRATA-G genotypes is described in [Chapter 7]. Imputation of CardiffCOGS followed the procedure used for the GWAS samples which relied on the 1000 Genomes reference panel, while STRATA-G was imputed using the Michigan Imputation Server and the Haplotype Reference Consortium panel (McCarthy et al., 2016). Again, similar imputation accuracy has been shown for common allele frequencies

using these two reference panels (McCarthy et al., 2016). The same post-imputation filters were applied to both CardiffCOGS and STRATA-G genotypes (INFO > 0.8; Probability threshold > 0.9; Missingness < 5%; Hardy Weinberg Equilibrium P-value > 10^{-6}).

Association testing

A GWAS of the TRS and non-TRS samples was carried out separately using PLINK v1.9. For this, we followed the procedures outlined in Pardiñas et al. (2018) and Ripke et al. (2014), including the control for population stratification based on PCA (Peloso & Lunetta, 2011).

Comparing the TRS and non-TRS association studies

In order to generate association statistics that reflect TRS/non-TRS differences, we used the 'test for interaction' proposed by Altman and Bland (2003), which is analogous to a fixed-effect test for moderators in the meta-analytic setting (Kontopantelis, Sperrin, Mamas, & Buchan, 2018). This test allows us to calculate an estimate (z-score) of the difference between two odds-ratios (OR). We calculated this z-score, and its associated P-value, at each overlapping SNP position between the TRS and non-TRS GWAS. For consistency, in all of our tests we used TRS as the 'reference' GWAS, and thus positive z-scores reflect stronger SNP-associations in TRS compared to non-TRS individuals.

Estimating genetic correlation

Genetic correlations between the TRS/non-TRS difference summary statistics and other complex traits were computed using LD-Hub v1.4 (Zheng et al., 2017), based on the LD-Score regression (Bulik-Sullivan et al., 2015b) framework. We only performed this analysis for GWAS categories which had some *a priori* likelihood of being related to psychiatric disorders, based on previous research (Krapohl et al., 2016). The categories chosen were 'education', 'neurological', 'personality', 'psychiatric' and 'smoking'. As the intelligence GWAS results from Sniekers et al. (2017) were not available in the LD-Hub platform at the time we conducted this study, we computed their genetic correlation statistics used the publicly available summary statistics and LD-Score v1.0 (Bulik-Sullivan et al., 2015b). To keep these results compatible with those from LD-Hub, we performed all the processing of

the summary statistics in an equivalent manner, by restricting to SNPs from the HapMap study (International HapMap 3 Consortium, 2010) and using pre-computed LD scores (Bulik-Sullivan et al., 2015b). Multiple testing correction of these results was carried out using Benjamini and Hochberg (1995) false discovery rate (FDR) method, since most of the summary statistics tested have overlapping samples and thus yield partially dependant P-values.

Estimating genetic co-localisation

The summary statistics that showed a significant genetic correlation with TRS, after multiple testing correction, were examined in more detail using GWAS-pw v0.21 (Pickrell et al., 2016). This software implements, in a genome-wide fashion, the Bayesian test for variant co-localisation proposed by Giambartolomei et al. (2014). Briefly, for each LD-independent region of the genome (Berisa & Pickrell, 2016), the test assesses if two GWAS share an association indexed by a single genetic marker. Such markers, if existing, can help to highlight causal relationships between traits, or shared underlying biological routes (Pickrell et al., 2016). Prior to performing our tests, whenever multiple studies examined the same or closely related trait, we retained only the most recent one, which in our case also had the larger sample size. From the GWAS-pw results we extracted the SNPs that showed a co-localisation posterior probability > 80%, and annotated them using the Variant Effect Predictor (VEP) tool (McLaren et al., 2016) available at Ensembl.

PRS analysis

Polygenic risk scores were estimated in PRSice-2 (Euesden, Lewis, & O'Reilly, 2014), using the interaction analysis and several GWAS as training sets, after restricting SNPs on each summary statistics file to those with a MAF of 10% or higher. All SNPs within the long-range LD regions detected by Price et al. (2008) were also excluded. The CardiffCOGS and STRATA-G samples constituted the testing sets, and P-value thresholds for the scores were set at eight different intervals ($P < 10^{-6}$, $P < 10^{-5}$, $P < 10^{-4}$, $P < 0.001$, $P < 0.05$, $P < 0.5$, $P < 0.1$, $P < 1$). Association of the polygenic scores with the TRS case/control phenotype was calculated using logistic models with the first 5 genotypic principal components (PCs) as covariates, and any other PC significantly associated to TRS, in order to control for population stratification (Peloso & Lunetta, 2011). Nagelkerke's R^2

values for each logistic regression were calculated on the liability scale (Lee, Goddard, Wray, & Visscher, 2012), assuming a 30% population prevalence for TRS (Lally et al., 2016b). Given that information on the operational criteria checklist (OPCRIT; McGuffin et al., 1991) was available for CardiffCOGS patients, we also carried out our polygenic score analysis by ascertaining TRS status through a positive rating in OPCRIT item 89 (treatment resistance).

Results

TRS/non-TRS interaction analysis

One genome-wide significant SNP was highlighted by the interaction analysis: rs79780963 ($P = 5.26 \times 10^{-9}$), an intronic SNP in the NT5C2 gene (Figure 11). While the SNP did not show any other genome-wide significant SNPs in close proximity or LD, rs79780963 has been previously described as an eQTL in brain of the ARL3 ($P = 1.40 \times 10^{-4}$) and NT5C2 ($P = 6.20 \times 10^{-4}$) genes, according to the BRAINEAC database (Ramasamy et al., 2014). Both of those genes belong to a GWAS locus which has been recently highlighted as associated to schizophrenia in European (Ripke et al., 2014) and Asian (Yu et al., 2017) populations.

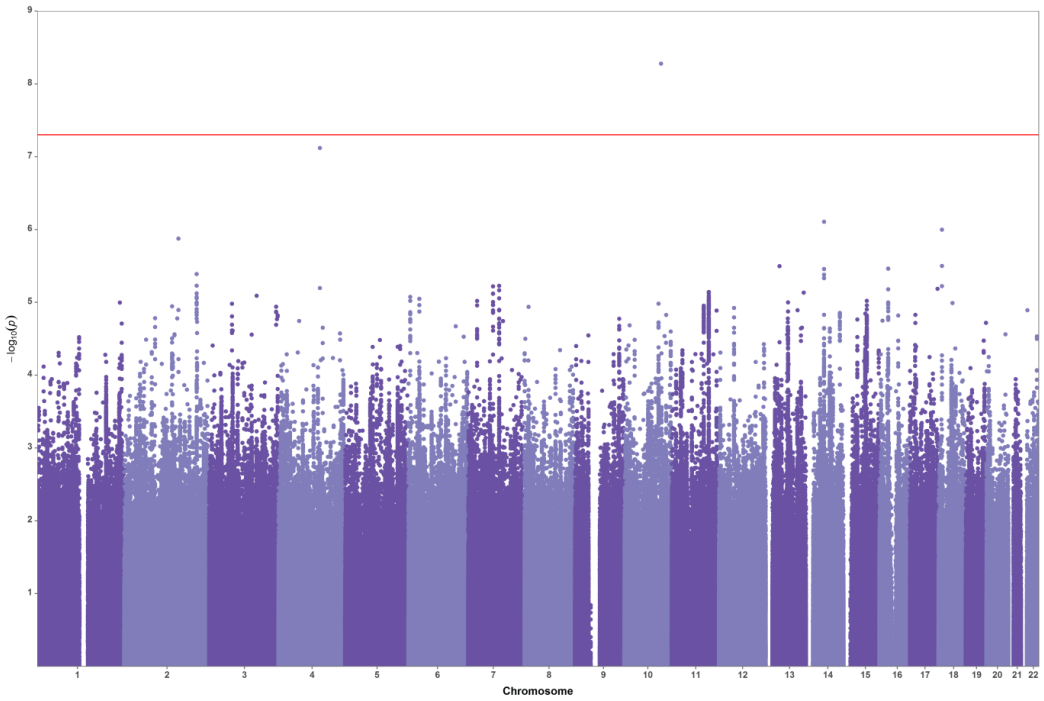


Figure 11. Manhattan plot of the TRS/non-TRS 'interaction' GWAS.

Genetic correlation and co-localisation

The analysis of 28 publicly available GWAS summary statistics showed that 10 of them displayed nominally significant ($P < 0.05$) genetic correlations, of which 6 survived multiple testing correction ($q < 0.05$; Figure 12). This set of summary statistics, which included educational attainment, adult intelligence and former smoking, showed similar genetic correlation (r_g) values with TRS. Given the small heritability detected for our TRS GWAS by the LD-Score method ($h^2_{SNP} = 0.014 \pm 0.005$), we note that these results might harbour some degree of statistical noise and thus should be interpreted cautiously.

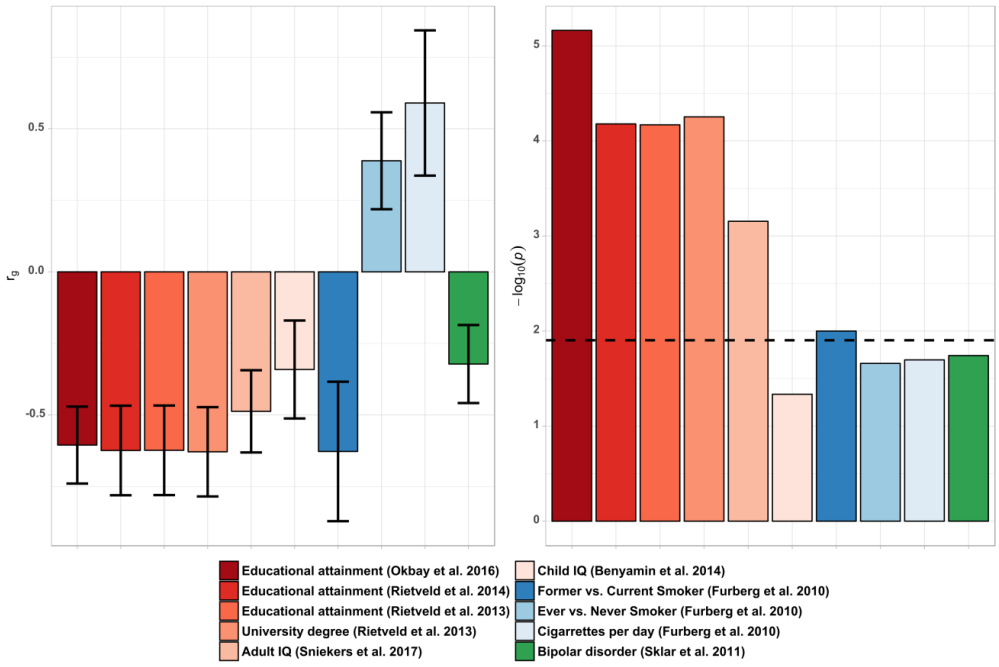


Figure 12. LDSR genetic correlation results of LD-HUB phenotypes with the TRS interaction GWAS. Left: Genetic correlation coefficients (r_g) shown with standard errors. Right: Genetic correlation p-values, the dashed line indicates an FDR-corrected significant level of $q = 0.05$.

A genetic co-localisation analysis highlighted that TRS also showed similar proportions of shared variants with the three phenotypes highlighted in the r_g analysis (Figure 13), though the trait was not as polygenic as education or intelligence, which have much larger numbers of associated variants. Focusing on particular genomic regions, we found 44 co-localised index SNPs between TRS,

educational attainment, adult intelligence and former smoking. All of these were non-coding variants, located either in introns or regulatory regions. The majority of these (39) were shared with educational attainment, with one of the regions (chr4: 118479918-119933512) shared also with intelligence via the SNP rs7664307. We note that this region, harbouring the NDST3 gene, has been highlighted before as genome-wide significant in schizophrenia and bipolar disorder in an Ashkenazi Jewish cohort (Lencz et al., 2013), though it did not achieve such significance in a follow-up study (Goes et al., 2015). As those studies do not mention that data on TRS or clozapine intake was available for their case samples, we cannot interpret the significance of this result in the context of a putative TRS-specific genetic architecture.

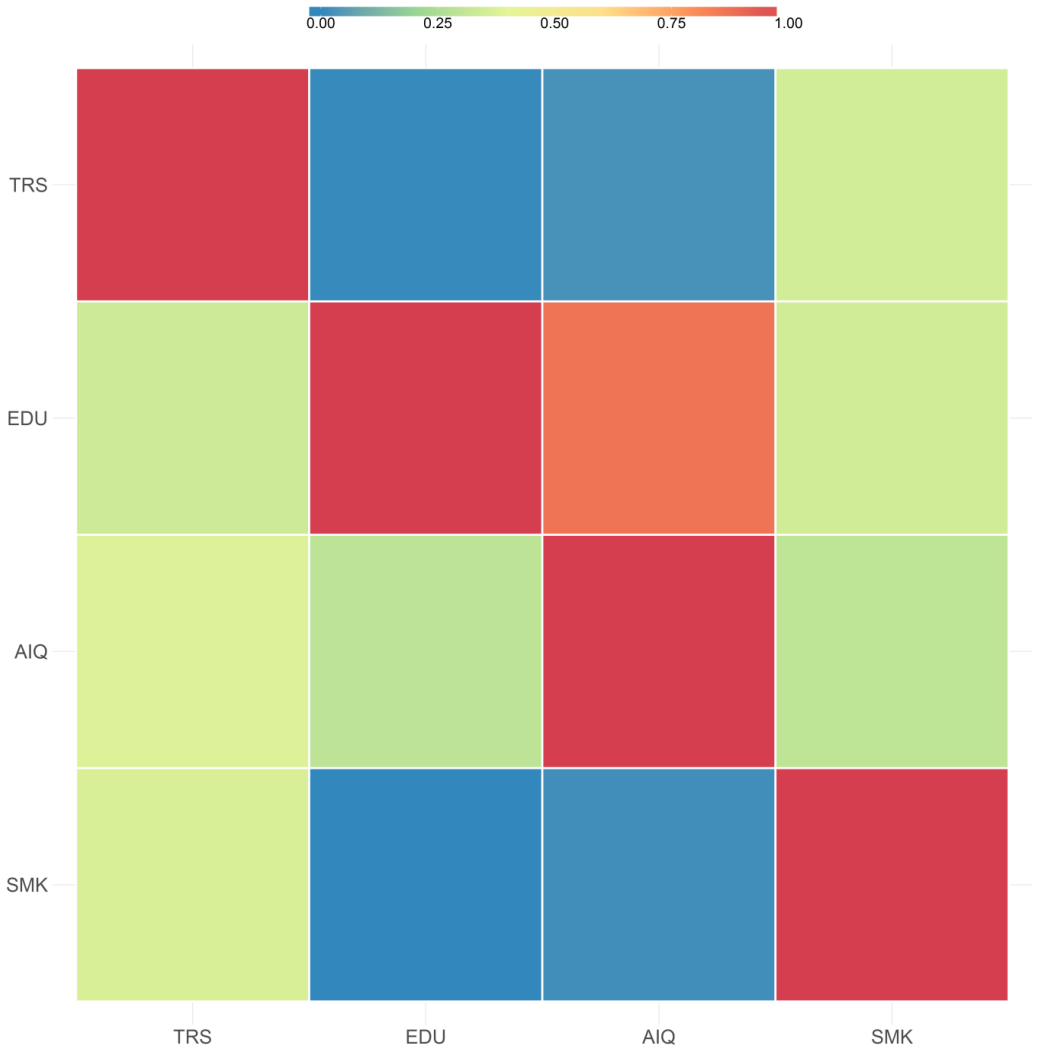


Figure 13. Pairwise matrix of GWAS-pw maximum a posteriori (MAP) estimates of shared variants between traits (TRS = Treatment-resistant schizophrenia; EDU

= Educational attainment; AIQ = Adult intelligence; SMK = Former vs. current smoking). Raw values show the proportion of SNPs that influence trait-x that also influence trait-y, or vice versa. Note that the different polygenic architectures of each trait make the matrix non-symmetrical.

Polygenic score-based prediction of TRS in independent samples

Our polygenic score analysis in the cross-sectional CardiffCOGS sample found that using the TRS GWAS as a training set could explain up to a 2.70% of the variance in the liability scale of treatment resistance, defined as a history of taking clozapine ($P = 1.42 \times 10^{-4}$). This was the strongest result obtained in our tests (Figure 14). Strikingly, a model based on GWAS results from the CLOZUK TRS sample also explained a significant proportion of the variance in the TRS phenotype (maximum $r^2 = 0.97\%$, $P = 0.022$), but no significant results were obtained with the much larger PGC2 non-TRS sample. A similar situation occurred when comparing the models based on intelligence and educational attainment, with only the former showing a significant association (maximum $r^2 = 1.03\%$, $P = 0.018$), despite the correlation in both phenotypes. Analysis of the first episode STRATA-G sample supported the results of CardiffCOGS, with significant variance explained for both the TRS GWAS (maximum $r^2 = 1.11\%$, $P = 0.046$) and CLOZUK (maximum $r^2 = 1.49\%$, $P = 0.016$), and consistent directions of effect (Figure 14).

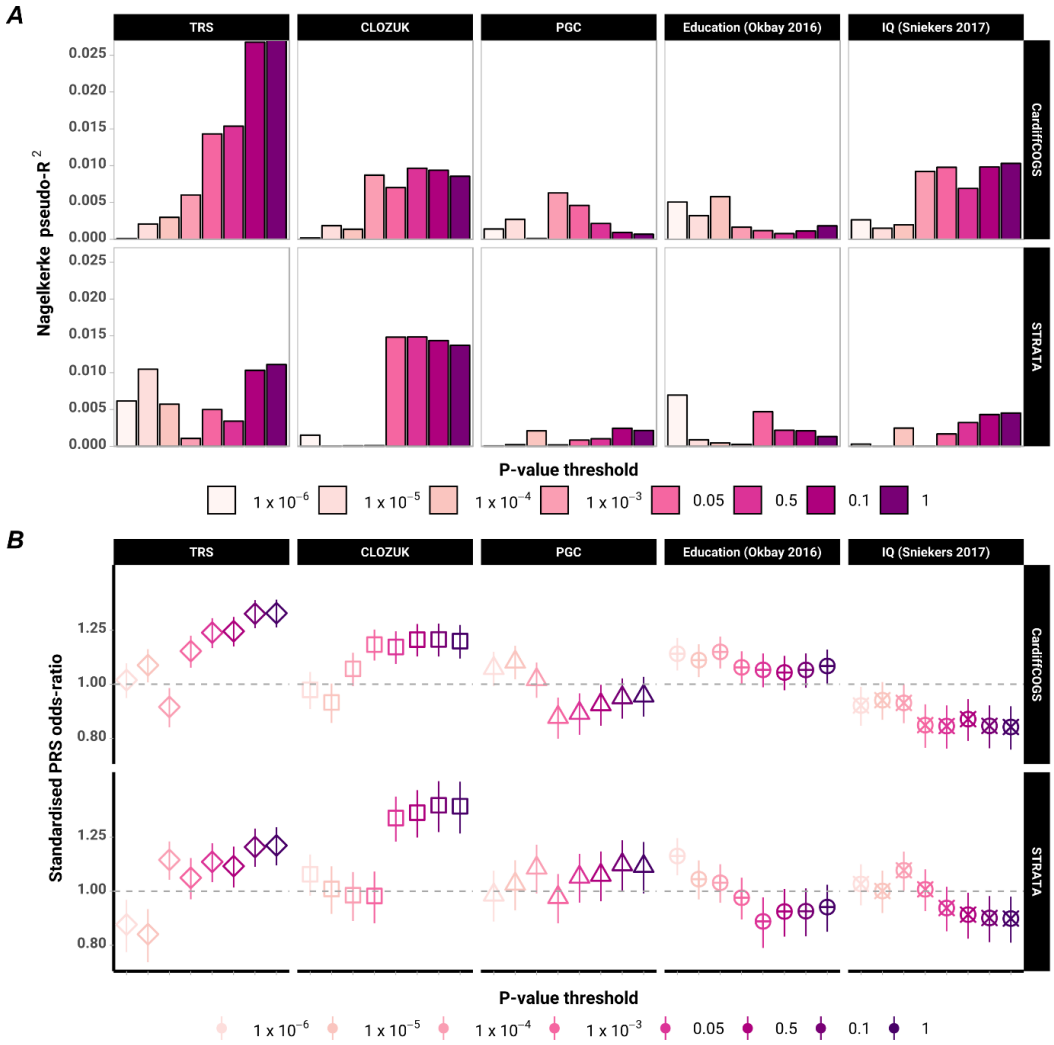


Figure 14. Polygenic score analysis of the CardiffCOGS and STRATA-G cohorts, using five different training sets. Explained variances on the liability scale (A) and effect sizes (B) are shown. The TRS outcome trait for this analysis was defined as a history of taking clozapine in CardiffCOGS (N_{TRS} = 341; N_{NONTRS} = 449) and treatment resistance in STRATA-G (N_{TRS} = 148; N_{NONTRS} = 911).

Finally, taking advantage of the deeper phenotyping of CardiffCOGS, we removing from the non-TRS subset those individuals positively rated for OPCRIT-89, which yielded a small increase in both significance and explained variance (Figure 15). Similar results were obtained when the TRS phenotype was directly defined as in the OPCRIT criteria (Figure 16).

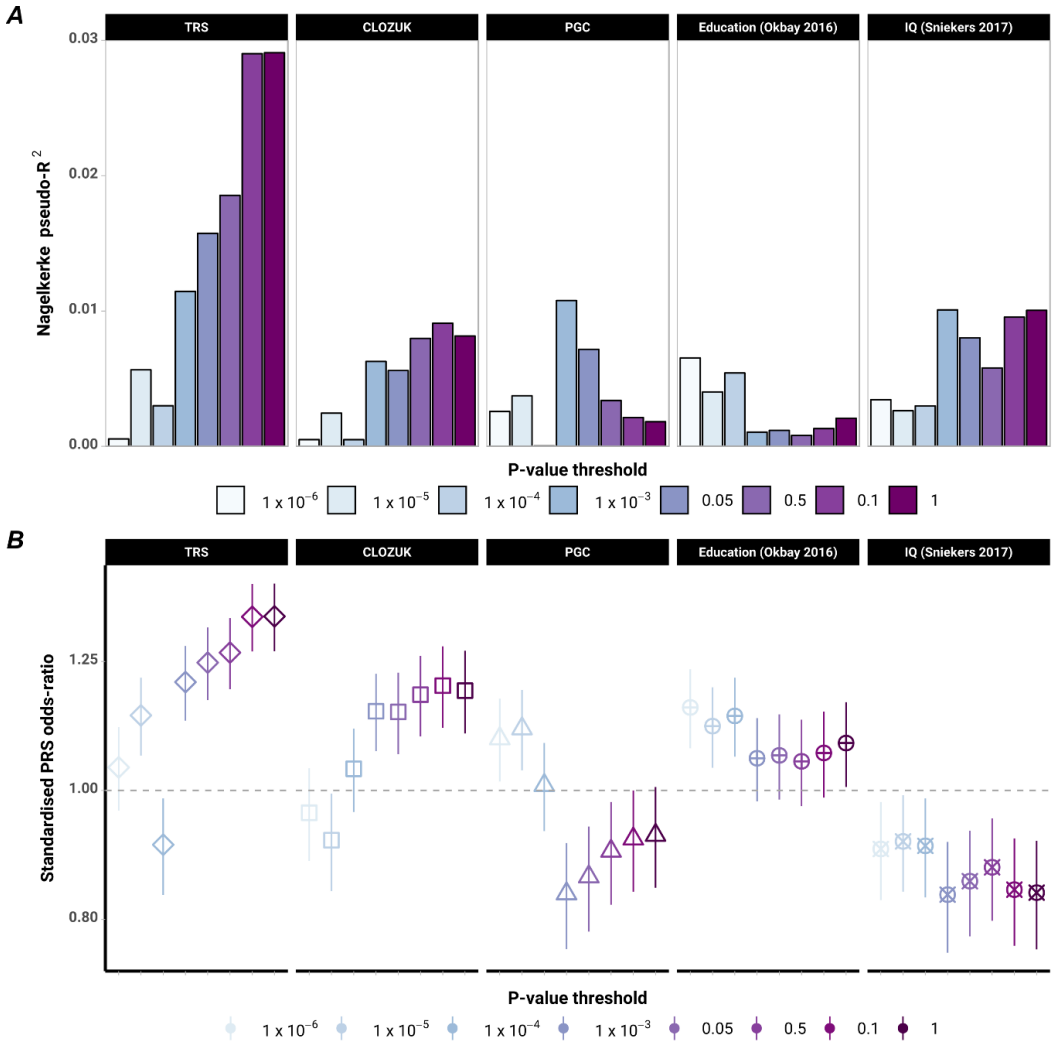


Figure 15. Polygenic score analysis of the CardiffCOGS cohort, using five different training sets. Explained variances on the liability scale (A) and effect sizes (B) are shown. The TRS outcome trait in this analysis was defined as a history of taking clozapine. Non-TRS individuals with positive OPCRIT-89 rating have been excluded from this analysis (N_{TRS} = 341; N_{NONTRS} = 386).

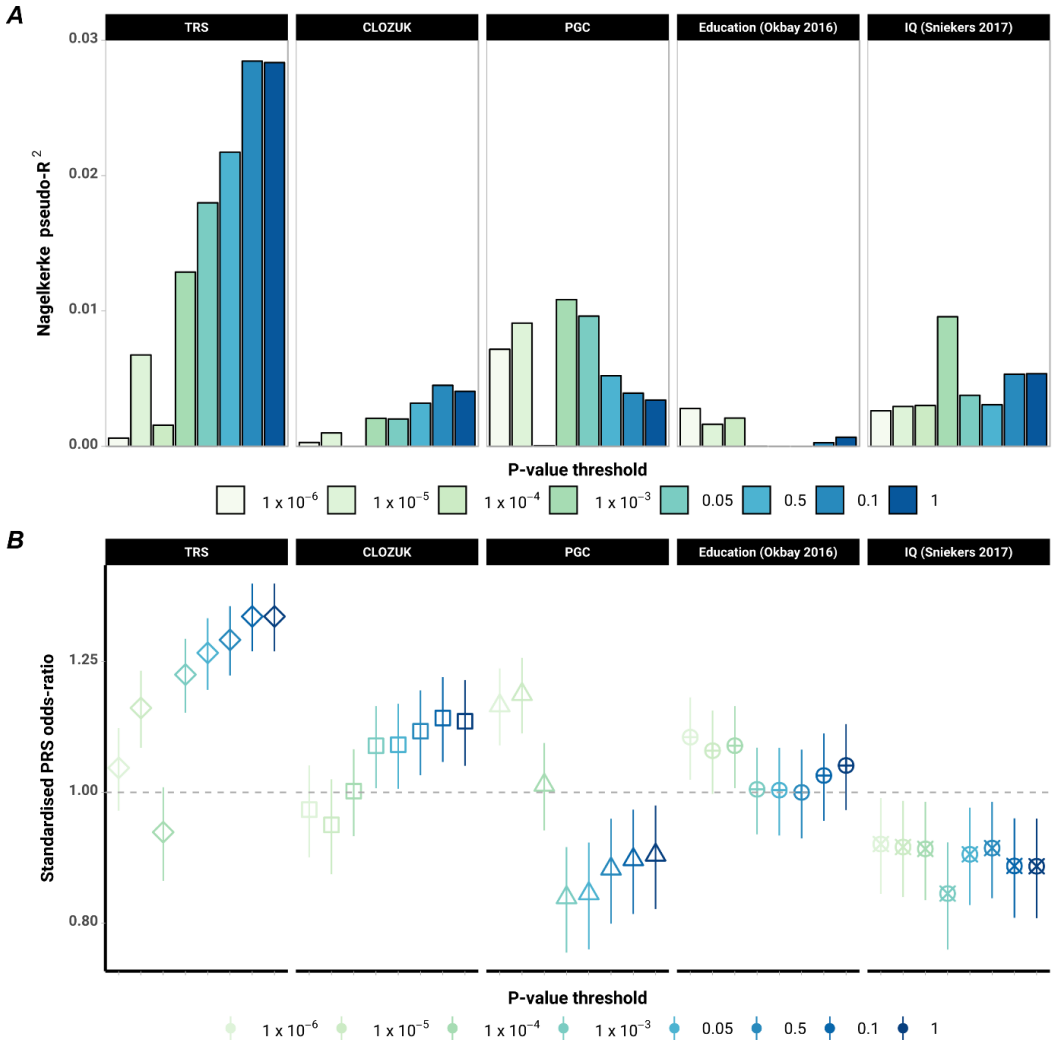


Figure 16. Polygenic score analysis of the CardiffCOGS cohort, using five different training sets. Explained variances on the liability scale (A) and effect sizes (B) are shown. The TRS outcome trait in this analysis was defined as a positive rating in the OPCRIT-89 item (N_{TRS} = 411; N_{NONTRS} = 350).

Discussion

The possible genetic basis of TRS has been a widely debated topic, and this study supports that a backbone of common variants is associated with this condition, and that these variants are not necessarily associated with other forms of schizophrenia. On the other hand, some of them are in common with traits that are both functionally correlated with TRS in previous research and genetically correlated in our analysis, such as proxies of cognition or smoking behaviour. These correlations are shown to be in the expected direction, with the group of cognitive phenotypes

negatively correlated to TRS, while the likelihood and amount of smoking are positively correlated. This recapitulates the findings from cross-sectional studies showing that TRS is associated with frequent smoking behaviour (Kennedy et al., 2014) and poorer cognitive performance (de Bartolomeis et al., 2013; Frydecka et al., 2016). However, while these results indicate polygenic commonalities between these traits, they cannot be interpreted as proof of causality in either direction, and they do not account for confounders such as socio-demographic indicators or the effect of antipsychotic medication.

Implications

In supporting the existence of a common, heritable genetic signal for TRS, our study adds a new layer of evidence to the discussion of whether TRS is "categorically distinct from treatment-responsive schizophrenia" (Gillespie et al., 2017). It also shows that the heterogeneous results previously obtained when attempting to predict TRS via schizophrenia polygenic scores might have been dependant on the existence of TRS individuals in the 'generic' schizophrenia training sample (Frank et al., 2015a), as well as in the classification strategy used to define TRS individuals in the target sample (Wimberley et al., 2017). While our study is not powered to take advantage of more sophisticated statistical methods, such as tissue-specific enrichment analyses or transcriptome-wide association studies (Gusev et al., 2016), it justifies the approach of performing direct GWAS in TRS and non-TRS individuals, in order to delve deeper into the TRS genetic architecture.

Limitations

A potential methodological concern is that the test used to perform the TRS interaction analysis is only sensitive to relatively large differences in effect size between studies (Stone-Romero & Anderson, 1994), and does not address potential between-study heterogeneity. While this could translate into residual inflation in the summary statistics, we do not detect this to be sizeable with internal quality-control metrics ($\lambda_{GC} = 1.055$, LD-Score intercept = 1.031 ± 0.007). Also, similar to what has been argued in GWAS of heterogeneous phenotypes (Niemi et al., 2018), the existence of significant genetic correlations and polygenic association in independent samples reassures us that the bulk of common variants we are

detecting are implicated in processes that can be related to the aetiology of TRS. Finally, given that the majority of our analyses are based on European and UK-based samples, our conclusions might not be generalizable to non-European countries where the diagnosis and treatment pathways of TRS might be influenced by ethnic or cultural backgrounds (Wheeler, 2008).

Conclusions

To our knowledge, this is the first study to show that genetic differences between TRS and non-TRS individuals can be aggregated into a polygenic genetic architecture, and that the condition itself has a small but detectable heritability. These results highlight the usefulness of well-controlled clinical phenotype data in the large samples currently recruited by international consortia working in psychiatric genetics, and how the use of this data can inform on the biological basis of psychiatric diagnostic classifications.

14.5. Appendix E

Table 29. Number of participants with genotypes that survived quality control, stratified by treatment response and cohort.

	Imputation Batch 1			**Imputation Batch 2**			**All**		
Cohort	NTR (%)	TR (%)	Total	NTR (%)	TR (%)	Total	NTR (%)	TR (%)	Total
AESOP London	13 (59.09)	9 (40.91)	22	NA	NA	NA	13 (59.09)	9 (40.91)	22
Belfast	15 (62.50)	9 (37.50)	24	NA	NA	NA	9 (37.50)	24	15 (62.50)
Bologna	3 (100)	NA	3	31 (93.94)	2 (6.06)	33	15 (62.50)	9 (37.50)	36
GAP London	NA	NA	NA	129 (80.12)	32 (19.88)	161	129 (94.44)	32 (19.88)	161
Istanbul	9 (45.00)	11 (55.00)	20	17 (73.91)	6 (26.09)	23	26 (60.47)	17 (39.53)	43
Lausanne	111 (88.10)	15 (11.90)	126	NA	NA	NA	111 (88.10)	15 (11.90)	126
Oslo	117 (97.50)	3 (2.50)	120	NA	NA	NA	117 (97.50)	3 (2.50)	120
Santander	313 (87.43)	45 (12.57)	358	NA	NA	NA	313 (87.43)	45 (12.57)	358
UCL London	32 (91.43)	3 (8.57)	35	NA	NA	NA	32 (91.43)	3 (8.57)	35
Total	613 (86.58)	95 (13.42)	708	177 (81.57)	40 (18.43)	217	790 (85.41)	135 (14.59)	925

Table 30. Average age of psychosis onset stratified by treatment response and cohort.

	NTR		**TR**		**All**	
Cohort	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
AESOP London	13	36.76 (13.10)	9	28.31 (8.80)	22	33.30 (12.07)
Belfast	15	26.78 (10.15)	9	23.56 (13.27)	24	25.58 (11.25)
Bologna	3	28.71 (10.64)	0	NA	3	19.45 (4.36)
Bologna (Batch 2)	31	29.61 (10.67)	2	29.45 (14.14)	33	29.60 (10.63)
GAP London (Batch 2)	129	28.43 (8.52)	32	23.11 (6.19)	161	27.37 (8.37)
Istanbul	9	21.12 (5.13)	11	19.53 (3.22)	20	20.05 (3.71)
Istanbul (Batch 2)	17	20.82 (5.62)	6	21 (3.90)	23	20.87 (5.14)
Lausanne	111	23.98 (5.19)	15	23.30 (5.61)	126	23.90 (5.22)
Oslo	117	24.27 (7.84)	3	22.00 (1.00)	120	24.22 (7.75)
Santander	313	29.28 (9.38)	45	26.68 (7.55)	358	28.95 (9.20)
UCL London	32	23.72 (7.69)	3	19.33 (2.52)	35	23.34 (7.47)
Total	790	27.21 (8.93)	135	24.23 (7.50)	925	26.78 (8.79)

Table 31. Partial Spearman's correlation coefficients between age of psychosis onset and each polygenic risk score stratified by treatment response and cohort.

	TR		**NTR**		**All**	
	N	r	N	r	N	r
TR-PRS						
AESOP London	13	0.68	9	-0.71	22	0.05
Belfast	15	NA	9	NA	24	0.10
Bologna	3	1	0	NA	3	0.61
Bologna (Batch 2)	31	-0.11	2	-1.00	33	-0.17
GAP London (Batch 2)	129	-0.01	32	-0.33	161	-0.13
Istanbul	8	0.49	11	-0.14	19	0.50
Istanbul (Batch 2)	17	0.12	6	0.56	23	0.11
Lausanne	109	-0.07	15	0.35	124	-0.07
Oslo	117	-0.06	3	-0.28	120	0.06
Santander	306	-0.02	45	0.00	351	-0.00
UCL London	32	-0.19	3	-1.00	35	-0.25
CLOZ-PRS						
AESOP London	13	0.68	9	0.20	22	0.26
Belfast	15	-0.15	9	0.92	24	0.28
Bologna	3	0.50	0	NA	3	0.50
Bologna (Batch 2)	31	-0.18	2	-1.00	33	-0.16
GAP London (Batch 2)	129	-0.12	32	-0.05	161	-0.11
Istanbul	8	0.71	11	-0.05	19	0.42
Istanbul (Batch 2)	17	0.32	6	0.34	23	0.12
Lausanne	109	-0.05	15	-0.21	124	-0.10
Oslo	117	0.02	3	1.00	120	0.02
Santander	306	0.02	45	-0.25	351	-0.01
UCL London	32	-0.09	3	0.76	35	-0.14
PGC-PRS						
AESOP London	13	-0.32	9	0.49	22	0.25
Belfast	15	0.07	9	0.41	24	0.08
Bologna	3	-1.00	0	NA	3	-1.00
Bologna (Batch 2)	31	-0.13	2	-1.00	33	-0.05
GAP London (Batch 2)	129	-0.15	32	0.46	161	0.04
Istanbul	8	-0.39	11	-0.04	19	-0.19
Istanbul (Batch 2)	17	0.14	6	0.35	23	0.13
Lausanne	109	0.06	15	0.11	124	-0.01
Oslo	117	0.07	3	-0.00	120	-0.02
Santander	306	-0.02	45	-0.01	351	-0.02
UCL London	32	0.19	3	1.00	35	0.07
EA-PRS						
AESOP London	13	NA	9	NA	22	0.12
Belfast	15	-0.24	9	-0.79	24	0.18
Bologna	3	0.13	0	NA	3	0.61
Bologna (Batch 2)	31	-0.28	2	1.00	33	-0.30
GAP London (Batch 2)	129	-0.10	32	-0.07	161	-0.04

Istanbul	8	-0.23	11	0.08	19	-0.17
Istanbul (Batch 2)	17	0.25	6	-0.24	23	0.35
Lausanne	109	0.00	15	0.62	124	-0.04
Oslo	117	0.06	3	-0.72	120	0.04
Santander	306	0.03	45	0.02	351	0.01
UCL London	32	-0.13	3	0.72	35	0.14
IQ-PRS						
AESOP London	13	NA	9	NA	22	-0.32
Belfast	15	NA	9	NA	24	-0.17
Bologna	3	0.13	0	NA	3	1.00
Bologna (Batch 2)	31	0.12	2	1.00	33	-0.02
GAP London (Batch 2)	129	0.06	32	-0.23	161	0.02
Istanbul	8	-0.29	11	0.33	19	0.49
Istanbul (Batch 2)	17	0.44	6	-0.34	23	0.18
Lausanne	109	-0.15	15	0.47	124	0.04
Oslo	117	-0.10	3	0.28	120	-0.06
Santander	306	0.01	45	0.02	351	-0.01
UCL London	32	-0.21	3	0.72	35	0.12

Table 32. Additional information on the meta-analyses of the partial Spearman's correlation coefficients between age of psychosis onset and each polygenic risk score.

When meta-analysed, there was no significant correlation between age of onset and TR-PRS (r = -0.01, standard error = 0.04, 95%CI: -0.10, 0.07). In terms of heterogeneity, Tau^2 = 0.0045 (standard error = 0.01, P = 0.09) and I^2 = 26.79%.

When meta-analysed, there was no significant correlation between age of onset and CLOZ-PRS (r = -0.02, standard error = 0.03, 95% CI = -0.08 – 0.05). In terms of heterogeneity, Tau^2 = 0 (standard error = 0.004, P = 0.16) and I^2 = 0.02%.

When meta-analysed, there was no significant correlation between age of onset and PGC-PRS (r = 0.00, standard error = 0.03, 95% CI = -0.06 – 0.07). In terms of heterogeneity, Tau^2 = 0 (standard error = 0.004, P = 0.95) and I^2 = 0%.

When meta-analysed, there was no significant correlation between age of onset and EA-PRS (r = 0.00, standard error = 0.03, 95% CI = -0.06 – 0.07). In terms of heterogeneity, Tau^2 = 0 (standard error = 0.004, P = 0.33) and I^2 = 0.01%.

When meta-analysed, there was no significant correlation between age of onset and IQ-PRS (r = 0.01, standard error = 0.03, 95% CI = -0.06 – 0.07). In terms of heterogeneity, Tau^2 = 0 (standard error = 0.004, P = 0.18) and I^2 = 0.02%.

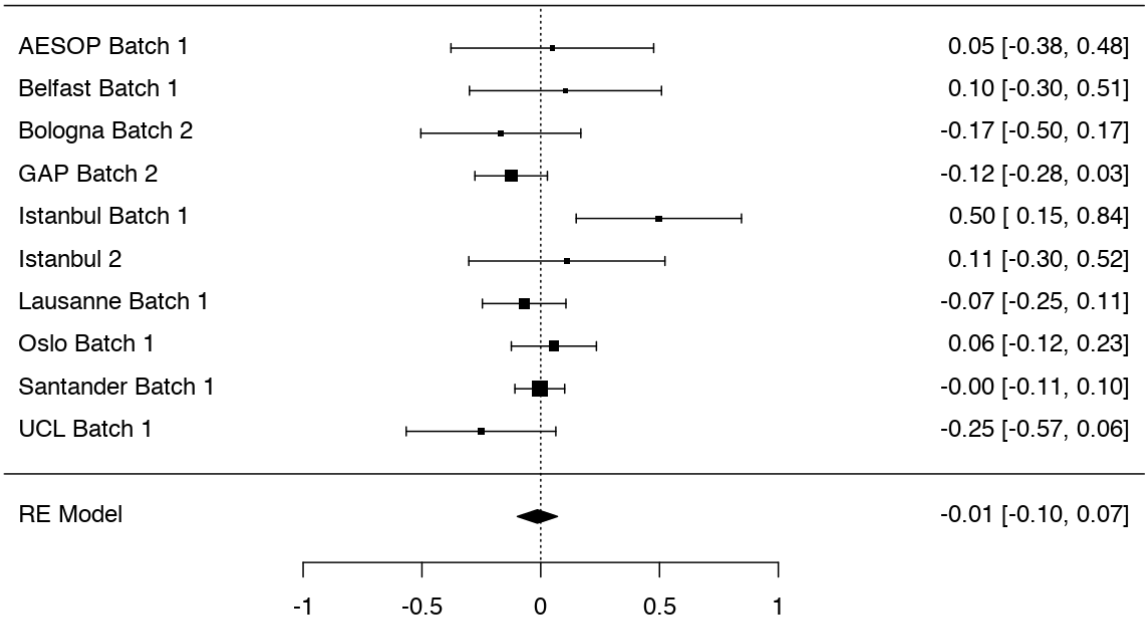


Figure 17. Forest plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for treatment resistance.

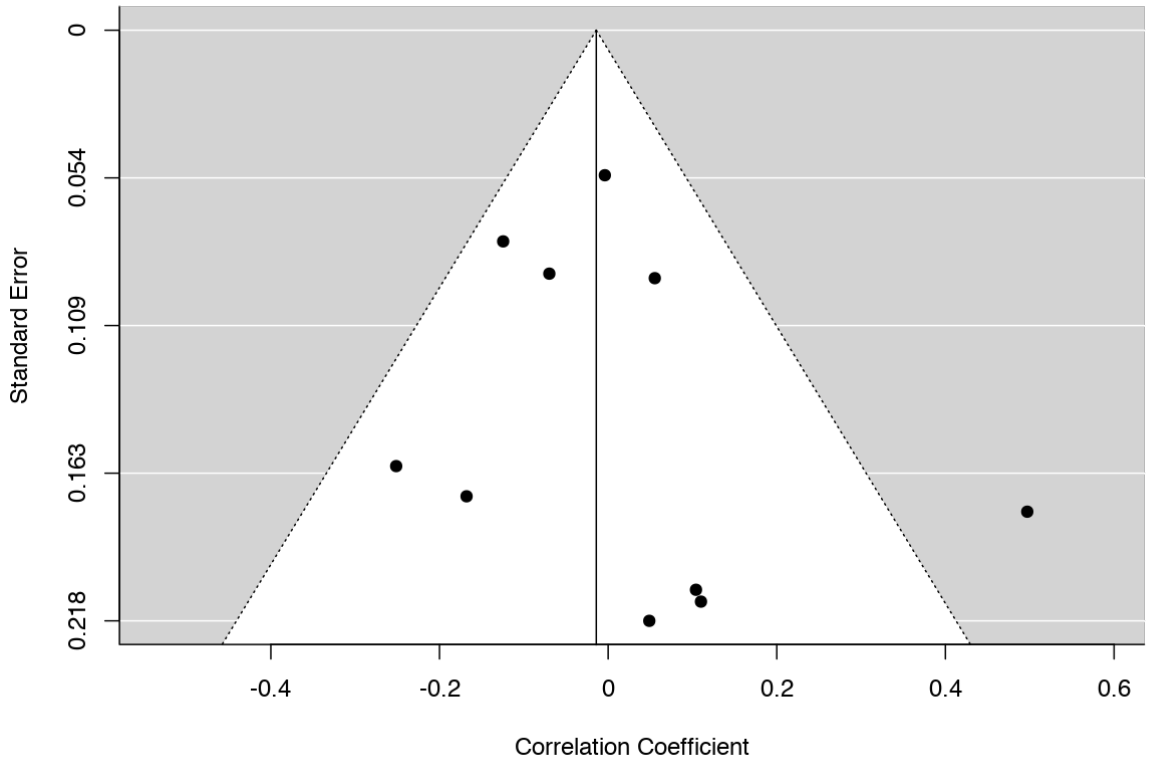


Figure 18. Funnel plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for treatment resistance.

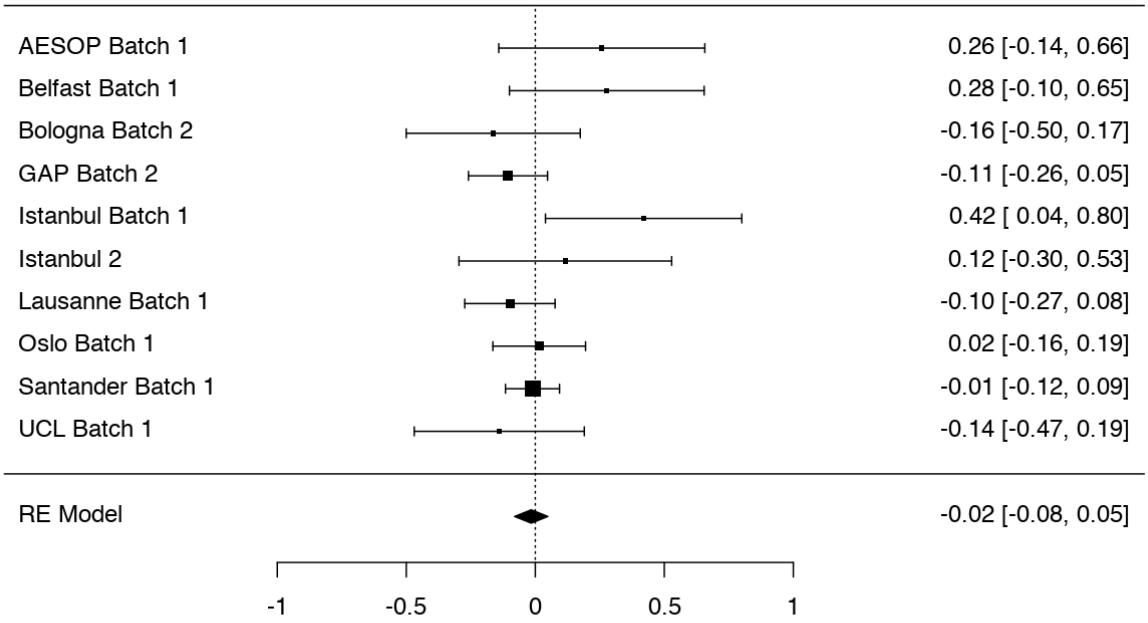


Figure 19. Forest plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for clozapine (clozapine vs healthy controls).

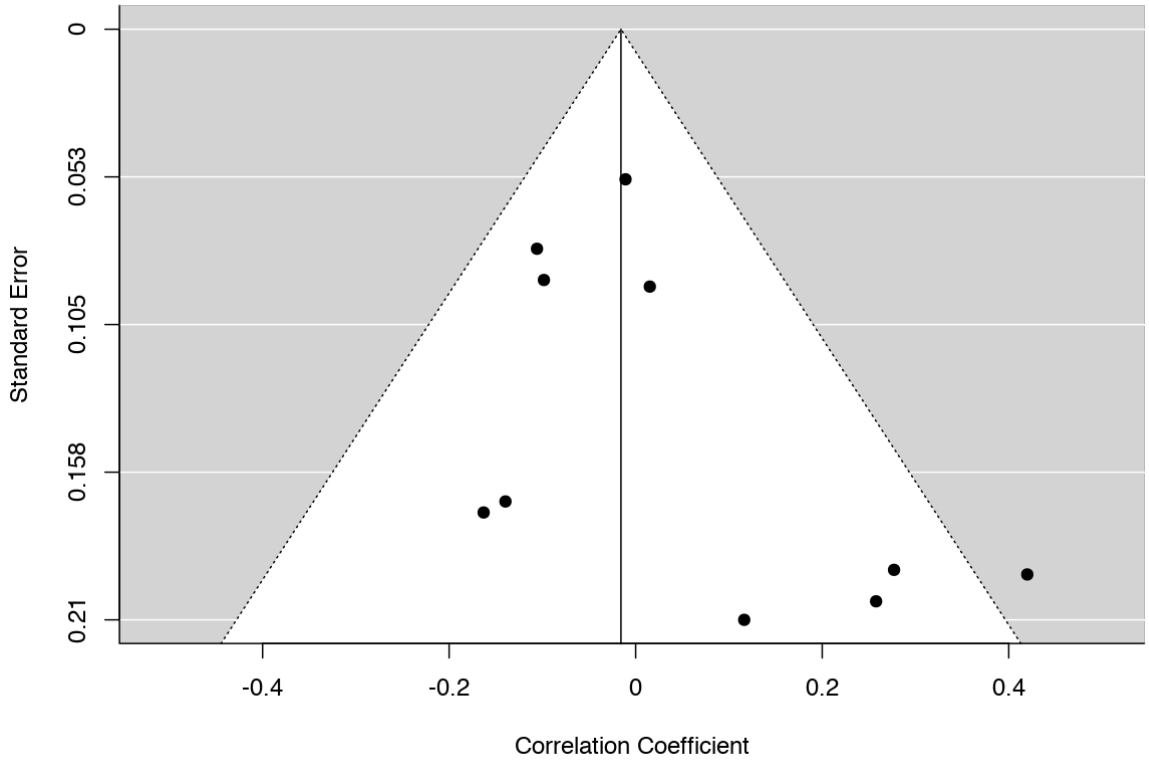


Figure 20. Funnel plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for clozapine (clozapine vs healthy controls).

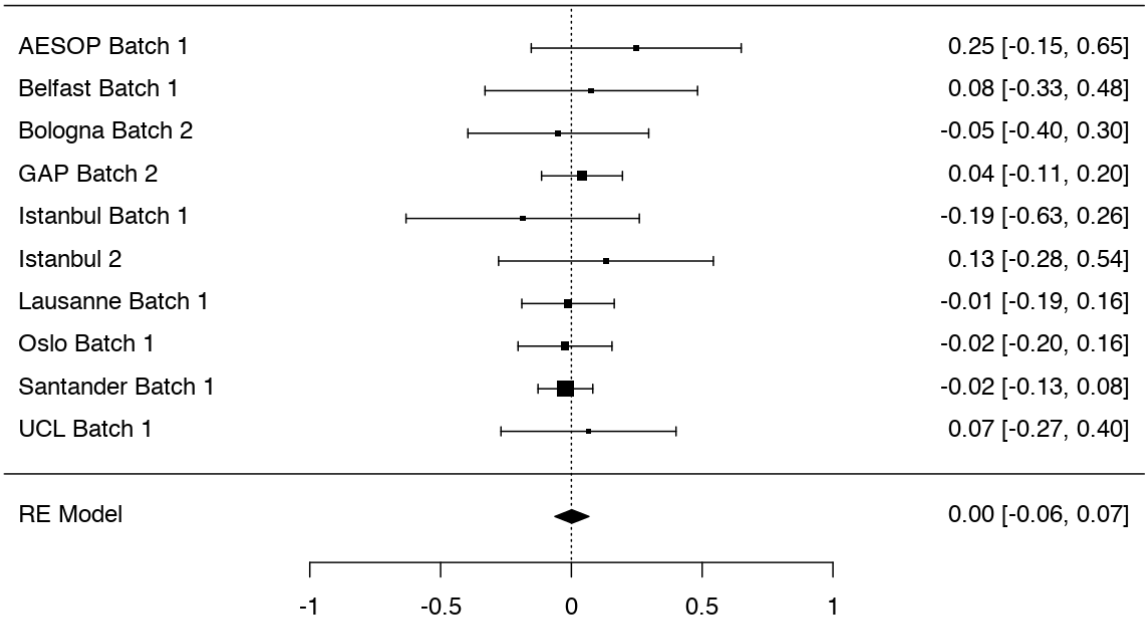


Figure 21. Forest plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for non-treatment resistant schizophrenia (non-treatment resistant schizophrenia vs healthy controls).

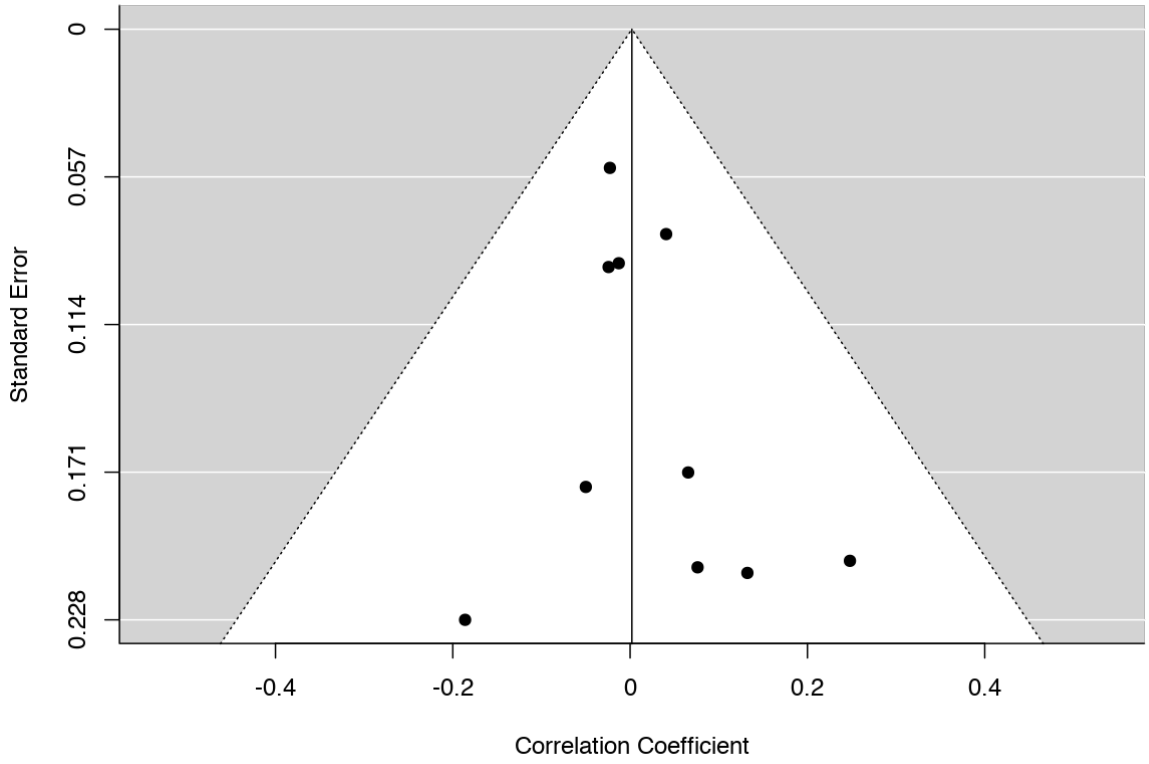


Figure 22. Funnel plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for non-treatment resistant schizophrenia (non-treatment resistant schizophrenia vs healthy controls).

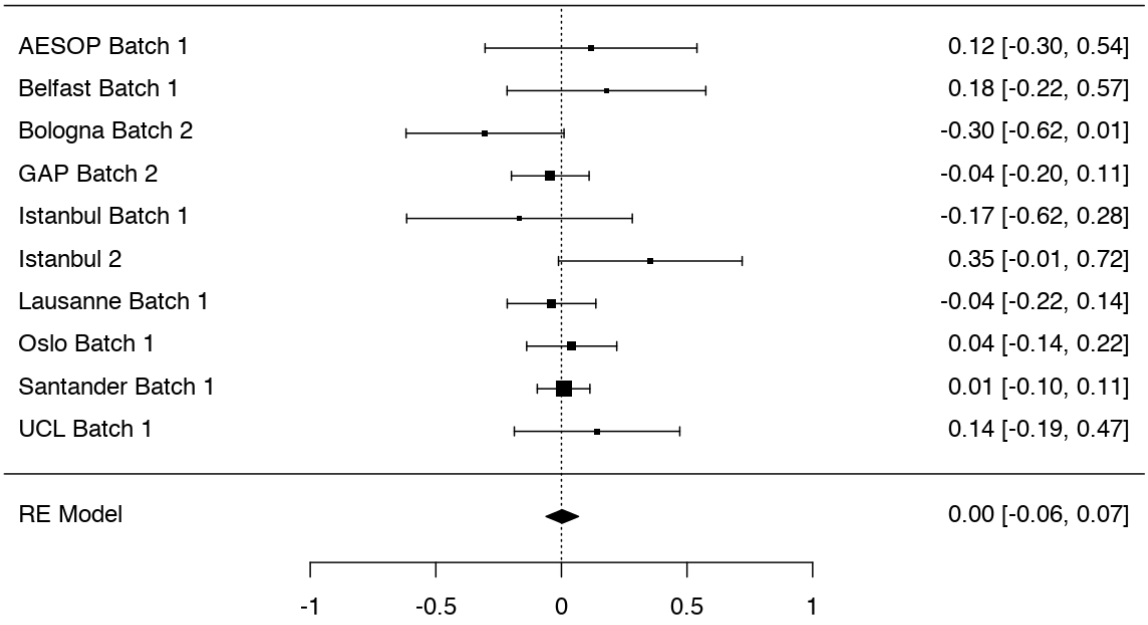


Figure 23. Forest plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for educational attainment.

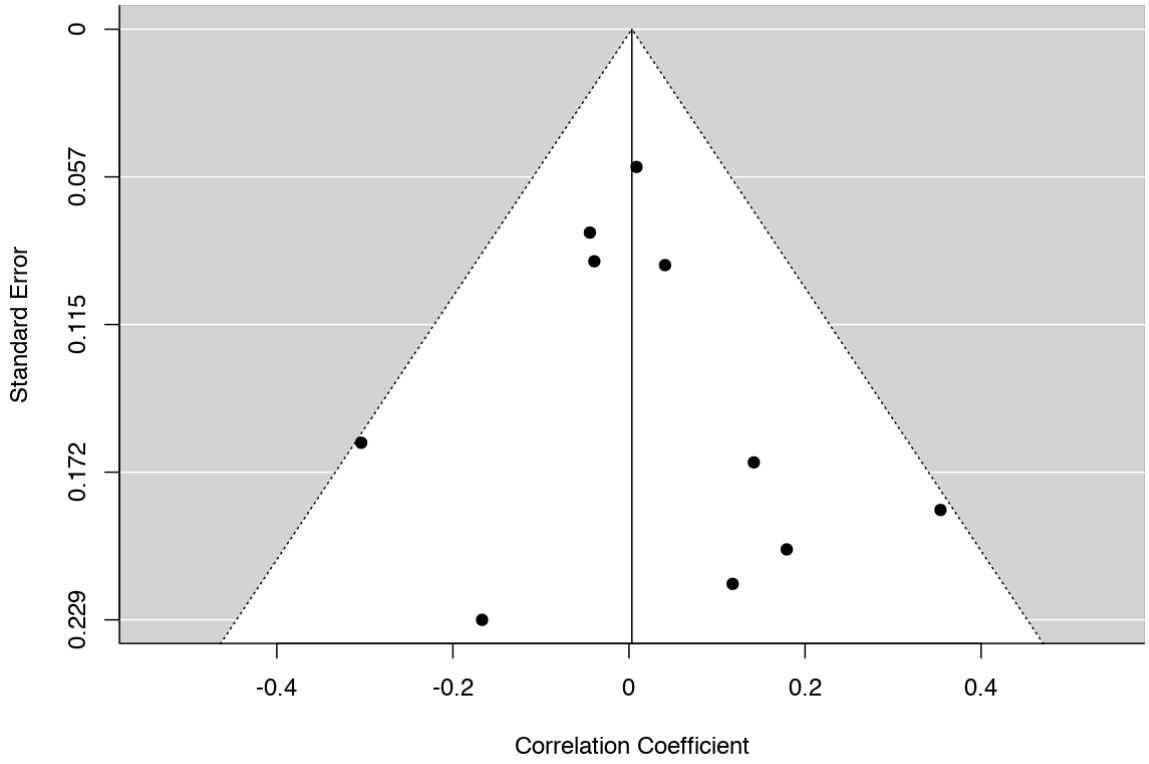


Figure 24. Funnel plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for educational attainment.

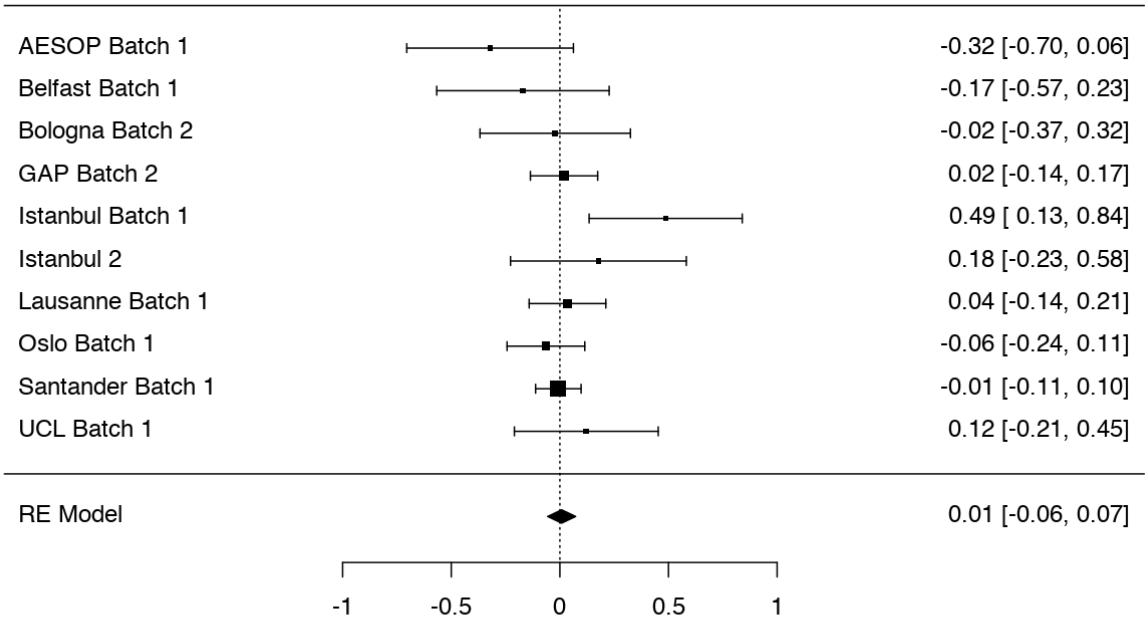


Figure 25. Forest plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for IQ.

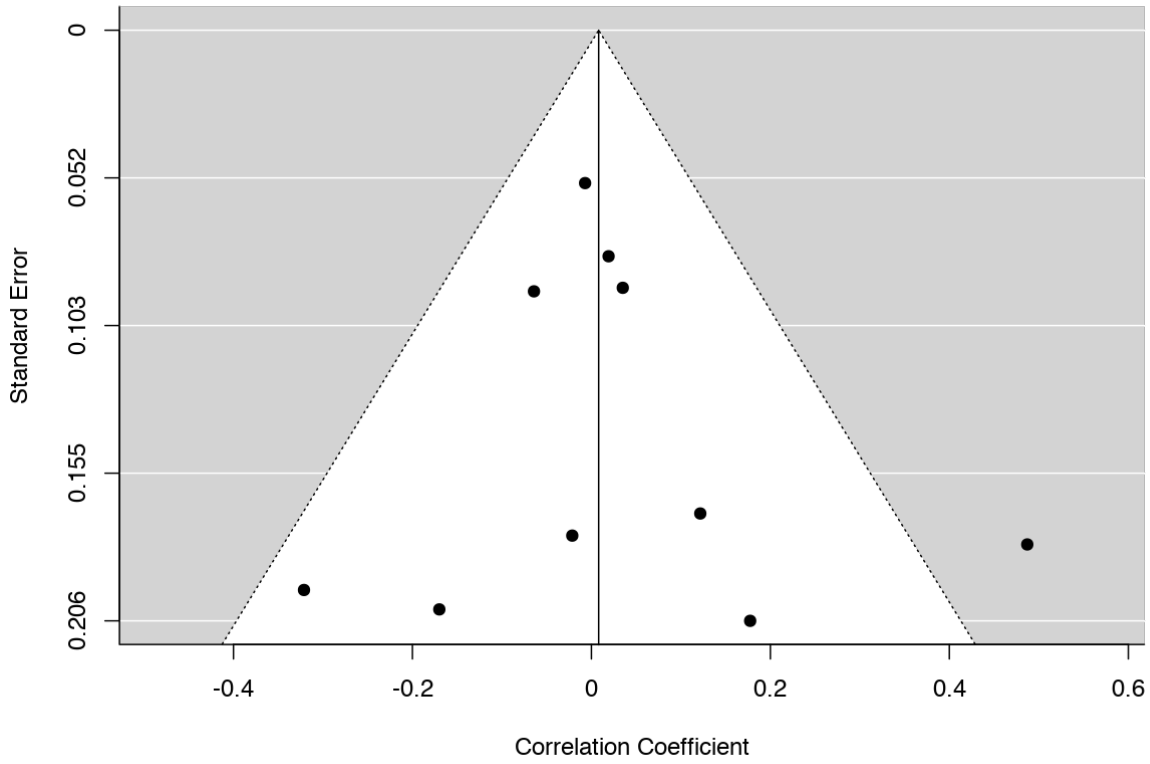


Figure 26. Funnel plot of partial Spearman's correlation between age of psychosis onset and the polygenic risk score for IQ.

Table 33. Average polygenic risk scores (calculated using LDpred-inf) stratified by treatment response and cohort.

	NTR		**TR (SD)**		**All**	
Cohort	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
TR-PRS						
AESOP London	13	-0.96 (1.96)	9	0.15 (1.21)	22	-0.50 (1.75)
Belfast	15	0.40 (2.90)	9	-0.76 (1.46)	24	-0.03 (2.49)
Bologna	3	-0.45 (0.82)	0	NA	3	-0.45 (0.81)
Bologna (Batch 2)	31	0.84 (2.83)	2	-0.72 (0.14)	33	0.75 (2.76)
GAP London (Batch 2)	129	-0.19 (2.07)	32	0.51 (2.34)	161	-0.05 (2.14)
Istanbul	8	0.83 (2.95)	11	-0.52 (1.56)	19	0.05 (2.28)
Istanbul (Batch 2)	17	-1.04 (2.18)	6	0.09 (1.70)	23	-0.75 (2.10)
Lausanne	109	0.06 (2.03)	15	0.55 (2.37)	124	0.12 (2.07)
Oslo	117	-0.06 (1.83)	3	1.03 (0.51)	120	-0.04 (1.82)
Santander	306	0.02 (2.17)	45	-0.12 (2.09)	351	0.001 (2.16)
UCL London	32	0.04 (2.27)	3	0.04 (2.14)	35	0.04 (2.23)
CLOZ-PRS						
AESOP London	13	-4.64 (15.33)	9	-5.36 (16.96)	22	-4.93 (15.62)
Belfast	15	-2.69 (36.52)	9	14.95 (12.49)	24	3.92 (30.70)
Bologna	3	-1.61 (6.22)	0	NA	3	-1.61 (6.22)
Bologna (Batch 2)	31	4.27 (21.47)	2	-6.07 (0.42)	33	3.65 (20.94)
GAP London (Batch 2)	129	-1.00 (17.08)	32	2.46 (17.69)	161	-0.32 (17.20)
Istanbul	8	14.28 (23.30)	11	3.93 (15.93)	19	8.29 (19.48)
Istanbul (Batch 2)	17	-6.70 (19.38)	6	7.42 (15.89)	23	-3.02 (19.26)
Lausanne	109	-2.55 (21.47)	15	4.80 (27.90)	124	-1.66 (22.34)
Oslo	117	0.10 (19.71)	3	-0.02 (6.76)	120	0.10 (19.48)
Santander	306	-0.21 (18.65)	45	1.41 (20.39)	351	0.001 (18.86)
UCL London	32	1.21 (20.37)	3	5.54 (29.73)	35	1.59 (20.78)
PGC-PRS						
AESOP London	13	4.63 (15.37)	9	-7.78 (21.63)	22	-0.45 (18.76)
Belfast	15	-9.18 (24.36)	9	29.78 (18.69)	24	5.43 (29.22)
Bologna	3	3.15 (2.50)	0	NA	3	3.15 (2.50)
Bologna (Batch 2)	31	-3.93 (25.59)	2	-5.33 (5.43)	33	-4.02 (24.80)
GAP London (Batch 2)	129	0.70 (20.03)	32	-1.23 (18.08)	161	0.31 (19.62)
Istanbul	8	6.97 (17.83)	11	10.29 (23.12)	19	8.89 (20.58)
Istanbul (Batch 2)	17	2.65 (22.21)	6	6.22 (16.67)	23	3.58 (20.60)
Lausanne	109	-4.60 (19.21)	15	0.09 (24.40)	124	-4.03 (19.85)
Oslo	117	1.45 (23.01)	3	-12.68 (12.05)	120	1.10 (22.88)
Santander	306	-0.17 (21.60)	45	2.21 (21.73)	351	0.14 (21.60)
UCL London	32	0.24 (20.07)	3	4.52 (13.68)	35	0.61 (19.49)
EA-PRS						
AESOP London	13	0.05 (0.21)	9	-0.03 (0.30)	22	0.02 (0.25)
Belfast	15	-0.03 (0.26)	9	-0.15 (0.24)	24	-0.07 (0.26)
Bologna	3	-0.10 (0.34)	0	NA	3	-0.09 (0.34)
Bologna (Batch 2)	31	-0.004 (0.28)	2	-0.02 (0.10)	33	-0.01 (0.27)
GAP London (Batch 2)	129	0.01 (0.24)	32	-0.03 (0.29)	161	0.001 (0.25)

Istanbul	8	0.02 (0.35)	11	0.03 (0.22)	19	0.024 (0.27)
Istanbul (Batch 2)	17	-0.03 (0.21)	6	0.10 (0.35)	23	0.002 (0.25)
Lausanne	109	-0.002 (0.26)	15	0.01 (0.25)	124	-0.0004 (0.26)
Oslo	117	0.03 (0.28)	3	0.07 (0.08)	120	0.03 (0.28)
Santander	306	-0.001 (0.27)	45	-0.05 (0.26)	351	-0.01 (0.27)
UCL London	32	0.01 (0.30)	3	0.04 (0.17)	35	0.02 (0.29)
IQ-PRS						
AESOP London	13	13.73 (67.81)	9	-10.67 (21.57)	22	3.75 (54.36)
Belfast	15	-11.08 (45.79)	9	-22.64 (56.95)	24	-15.41 (49.36)
Bologna	3	17.76 (27.14)	0	NA	3	17.76 (27.14)
Bologna (Batch 2)	31	-10.06 (53.11)	2	12.59 (33.76)	33	-8.69 (52.06)
GAP London (Batch 2)	129	1.46 (40.32)	32	-3.78 (42.50)	161	0.42 (40.68)
Istanbul	8	1.02 (60.07)	11	-3.33 (49.76)	19	-1.50 (52.76)
Istanbul (Batch 2)	17	9.04 (35.08)	6	10.88 (66.58)	23	9.52 (43.62)
Lausanne	109	7.82 (41.07)	15	-1.12 (50.90)	124	6.73 (42.24)
Oslo	117	4.27 (43.68)	3	-35.44 (27.66)	120	3.28 (43.72)
Santander	306	-1.48 (47.03)	45	-8.23 (47.06)	351	-2.35 (47.02)
UCL London	32	-2.90 (57.52)	3	-16.52 (31.39)	35	-4.07 (55.58)

Table 34. Additional information on the meta-analyses of the means for each polygenic risk score between TR and NTR patients.

There was no difference in the TR-PRS between TR and NTR patients ($t(191) = -0.49$, $P = 0.60$). When meta-analysed, there was no significant difference between TR and NTR patients in terms of TR-PRS ($r = 0.12$, standard error = 0.34, 95%CI: -0.54, 0.79). In terms of heterogeneity, $Tau^2 = 0.68$ (standard error = 0.52, $P < 0.001$) and $I^2 = 68.51\%$.

There was no difference in the CLOZ-PRS between TR and NTR patients ($t(185) = -1.9$, $P = 0.06$). When meta-analysed, there was no significant difference between TR and NTR patients in CLOZ-PRS ($r = 1.11$, standard error = 2.52, 95%CI: -3.84, 6.06). In terms of heterogeneity, $Tau^2 = 25.78$ (standard error = 27.50, $P = 0.06$) and $I^2 = 47.17\%$.

There was no difference in the PGC-PRS between TR and NTR patients ($t(182) = -1.6$, $P = 0.1$). When meta-analysed, there was no significant difference between TR and NTR patients in PGC-PRS ($r = 4.96$, standard error = 6.32, 95%CI: -7.42, 17.35). In terms of heterogeneity, $Tau^2 = 214.93$ (standard error = 169.88, $P < 0.001$) and $I^2 = 67.80\%$.

There was no difference in the EA-PRS between TR and NTR patients ($t(187) = 1.1$, $P = 0.3$). When meta-analysed, there was no significant difference between TR and NTR patients in EA-PRS ($r = -0.01$, standard error = 10.40, 95%CI: -20.39, 20.378). In terms of heterogeneity, $Tau^2 = 0$ (standard error = 438.90, $P = 1.00$) and $I^2 = 0\%$.

There was no difference in the IQ-PRS between TR and NTR patients ($t(183) = 1.9$, $P = 0.06$). When meta-analysed, there was no significant difference between TR and NTR patients in IQ-PRS ($r = -0.17$, standard error = 0.10, 95%CI: -0.36, 0.22). In terms of heterogeneity, $Tau^2 = 0$ (standard error = 0.04, $P = 0.97$) and $I^2 = 0\%$.

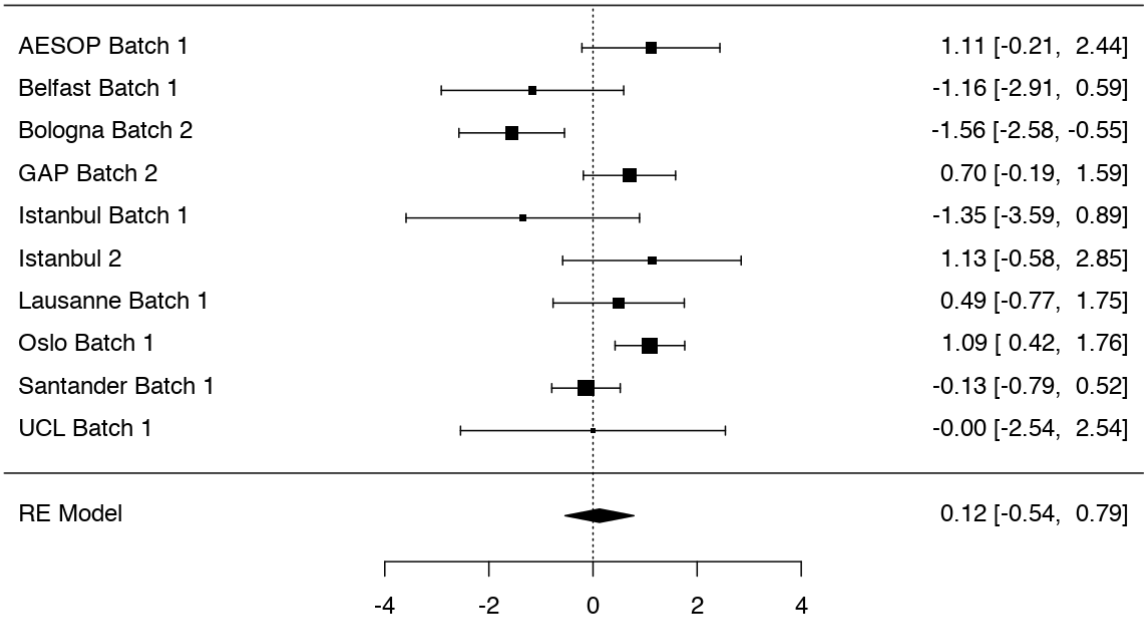


Figure 27. Forest plot of the raw mean difference in TR-PRS between treatment resistant and non-treatment resistant participants.

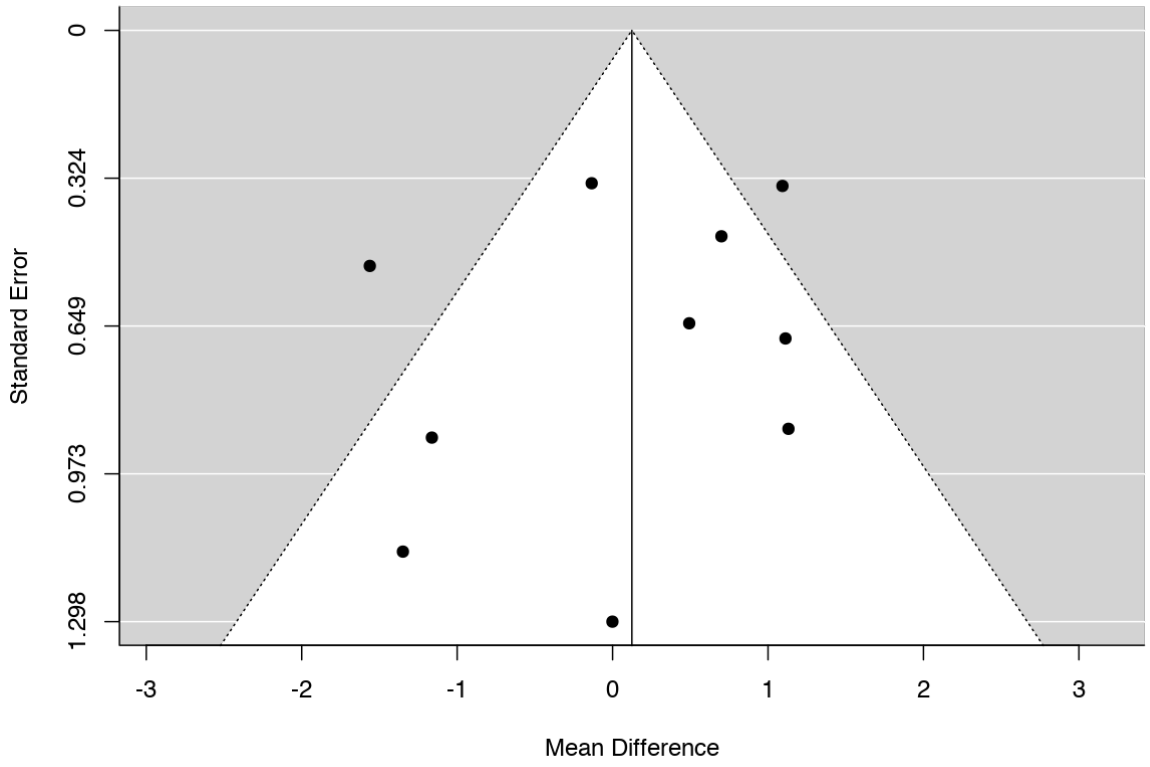


Figure 28. Funnel plot of the raw mean difference in TR-PRS between treatment resistant and non-treatment resistant participants.

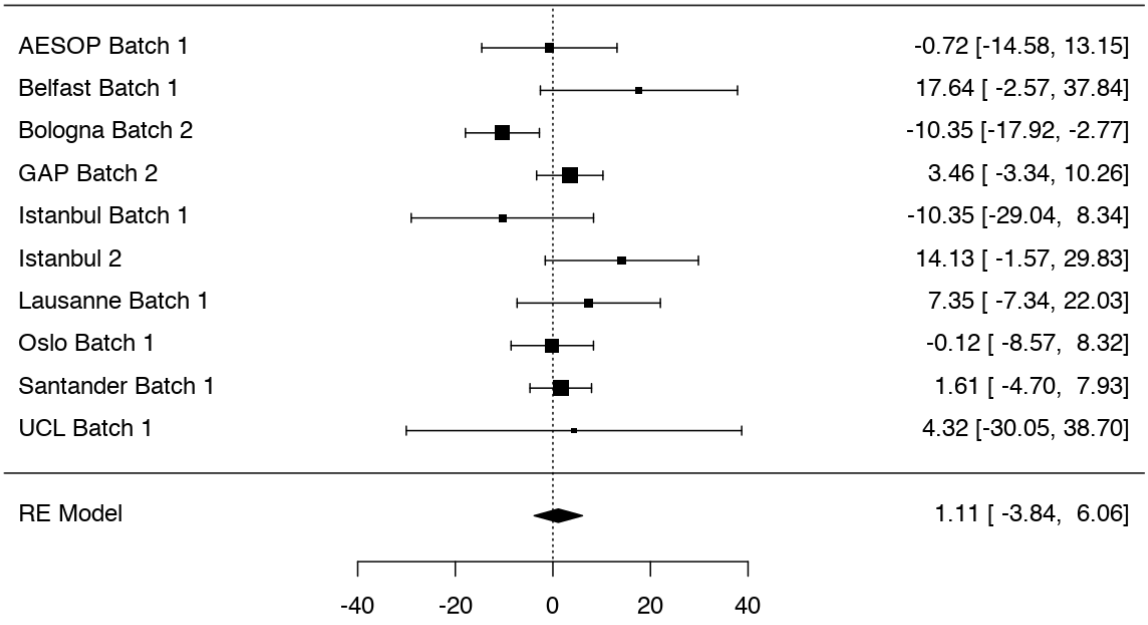


Figure 29. Forest plot of the raw mean difference in CLOZ-PRS between treatment resistant and non-treatment resistant participants.

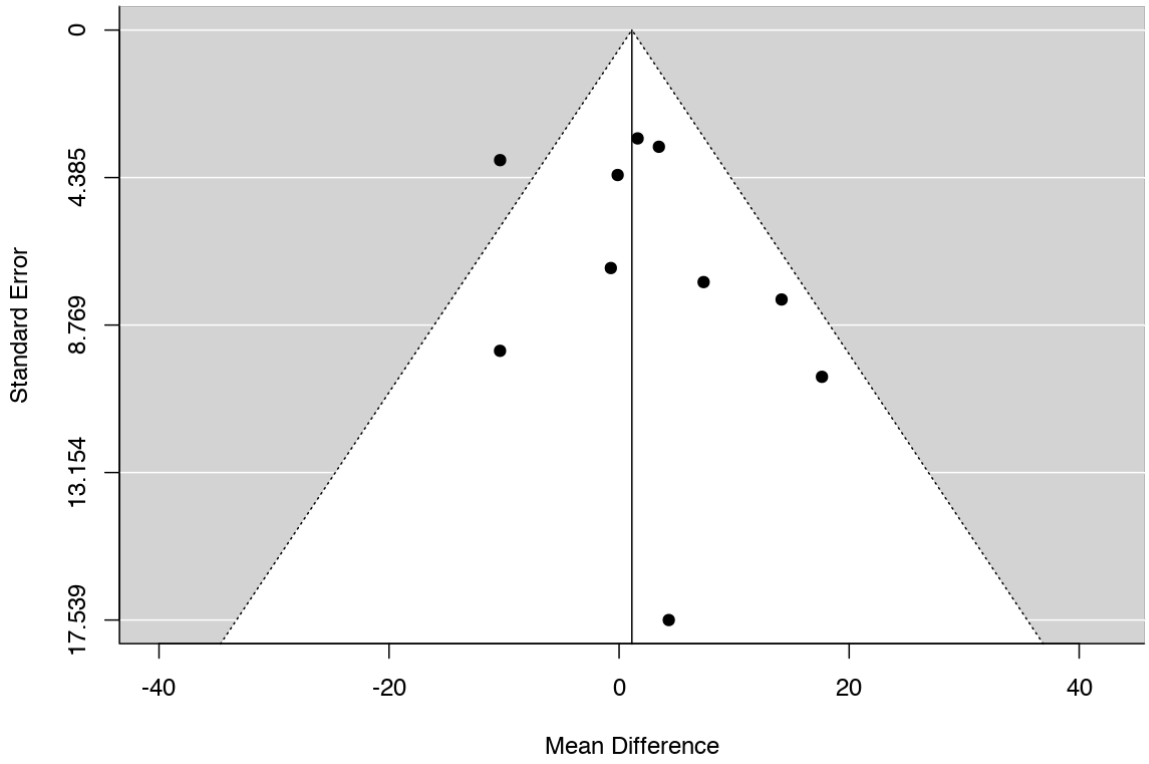


Figure 30. Funnel plot of the raw mean difference in CLOZ -PRS between treatment resistant and non-treatment resistant participants.

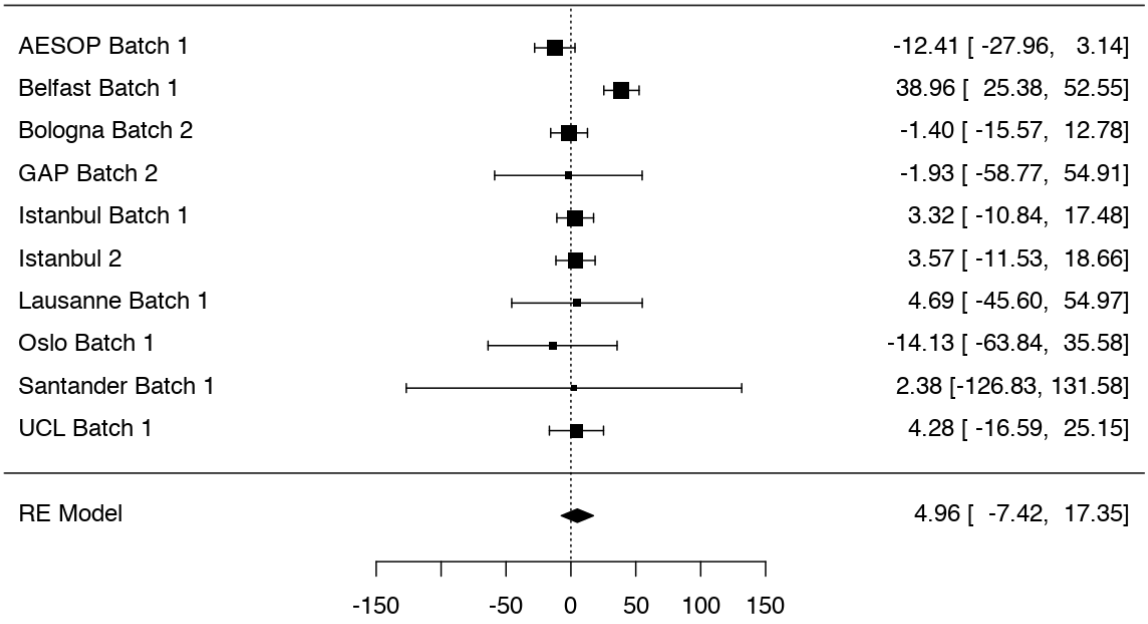


Figure 31. Forest plot of the raw mean difference in PGC-PRS between treatment resistant and non-treatment resistant participants.

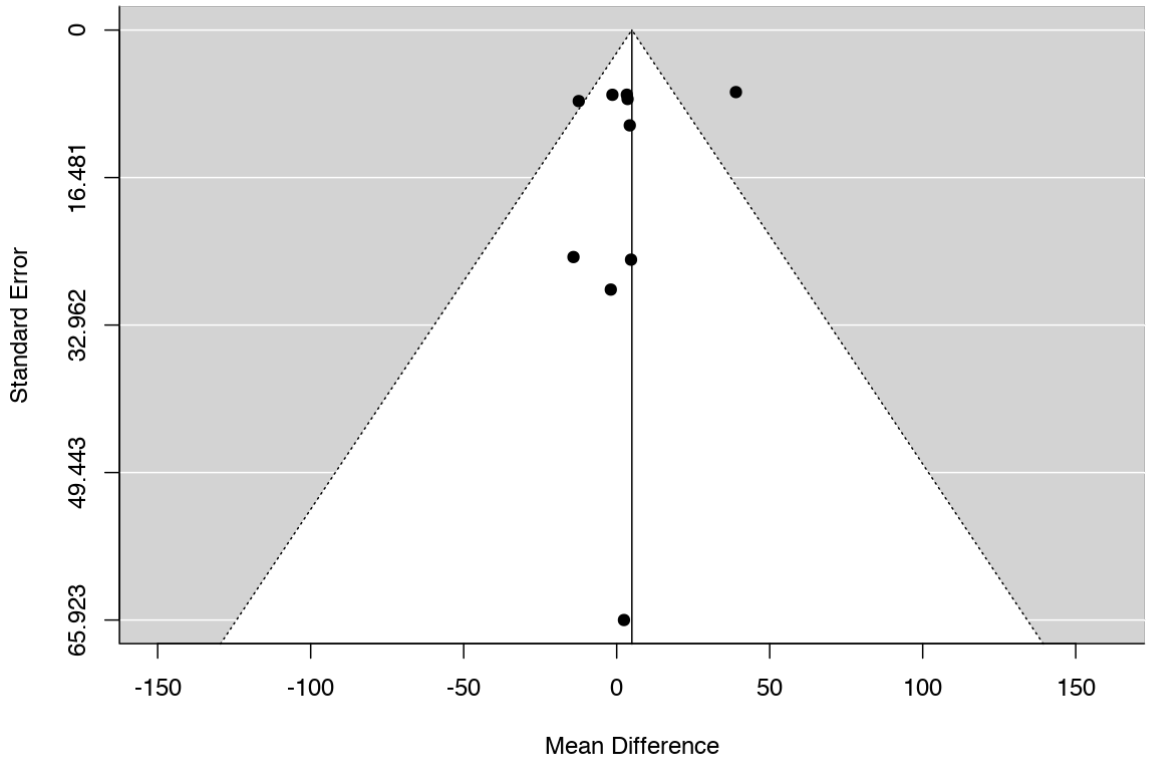


Figure 32. Funnel plot of the raw mean difference in PGC -PRS between treatment resistant and non-treatment resistant participants.

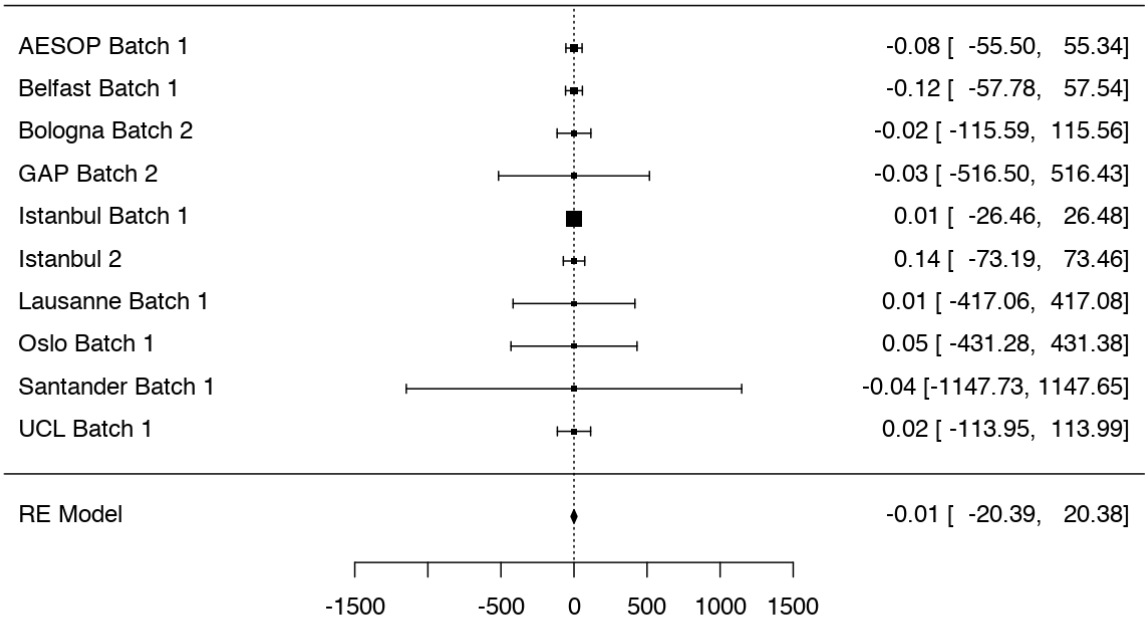


Figure 33. Forest plot of the raw mean difference in EA-PRS between treatment resistant and non-treatment resistant participants.

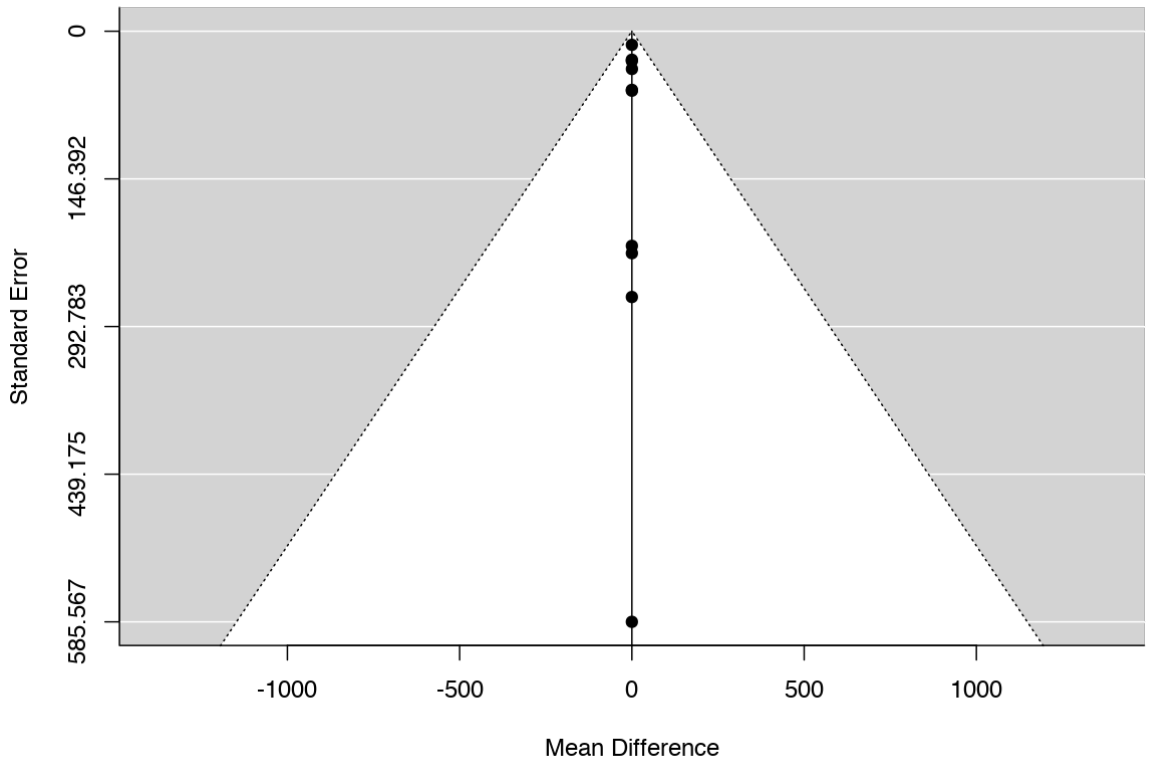


Figure 34. Funnel plot of the raw mean difference in EA -PRS between treatment resistant and non-treatment resistant participants.

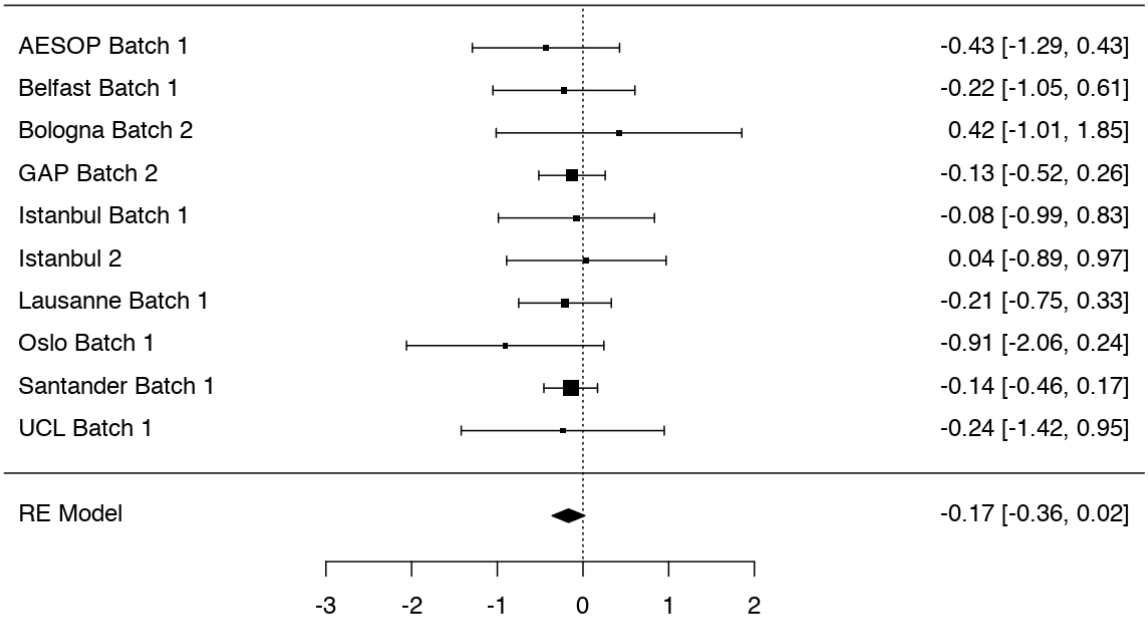


Figure 35. Forest plot of the raw mean difference in IQ-PRS between treatment resistant and non-treatment resistant participants.

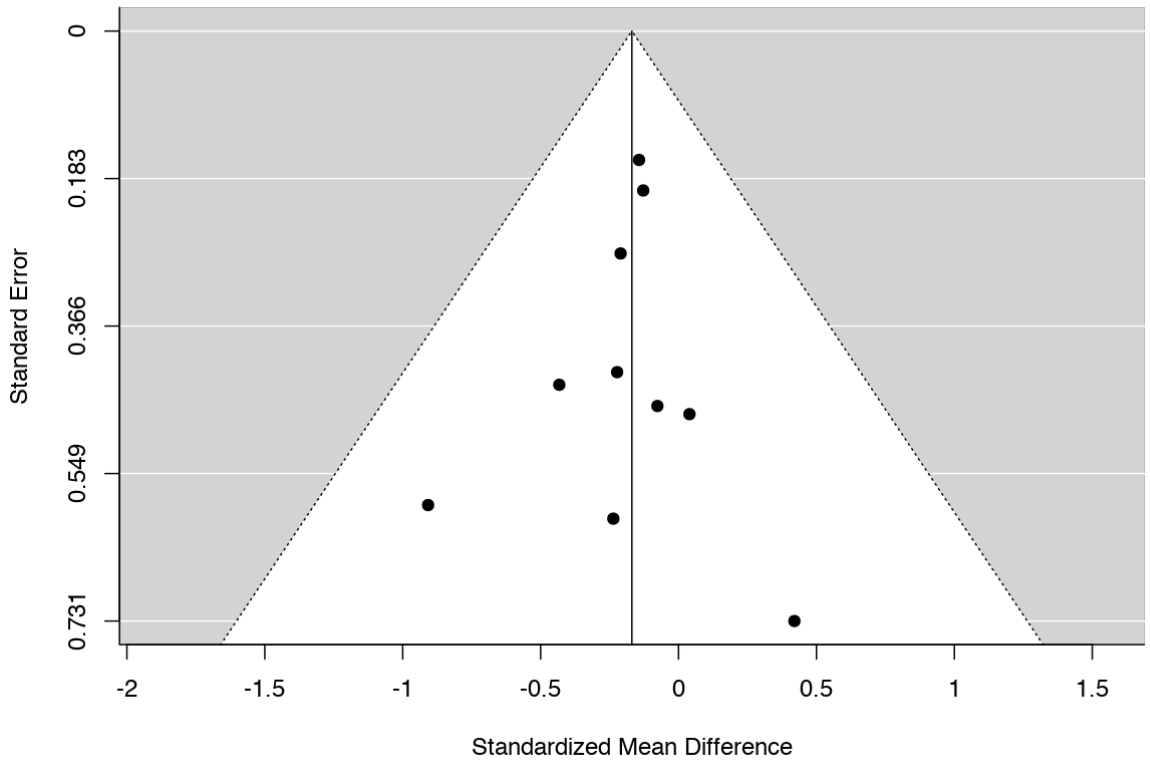


Figure 36. Funnel plot of the raw mean difference in IQ -PRS between treatment resistant and non-treatment resistant participants.

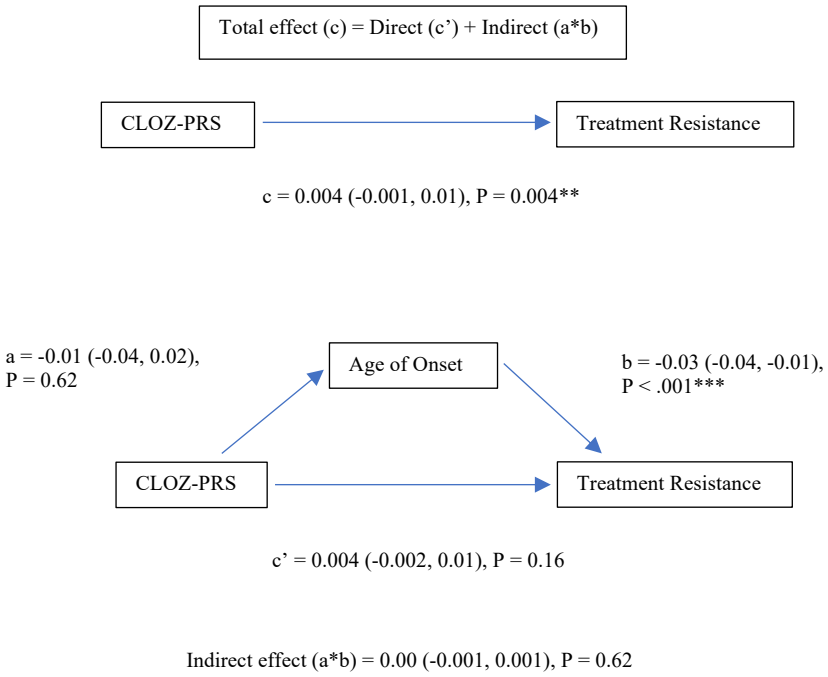


Figure 37. Output diagram showing the direct effect between the polygenic risk score for clozapine and treatment resistance, and the indirect effect with age of psychosis onset as a mediator.

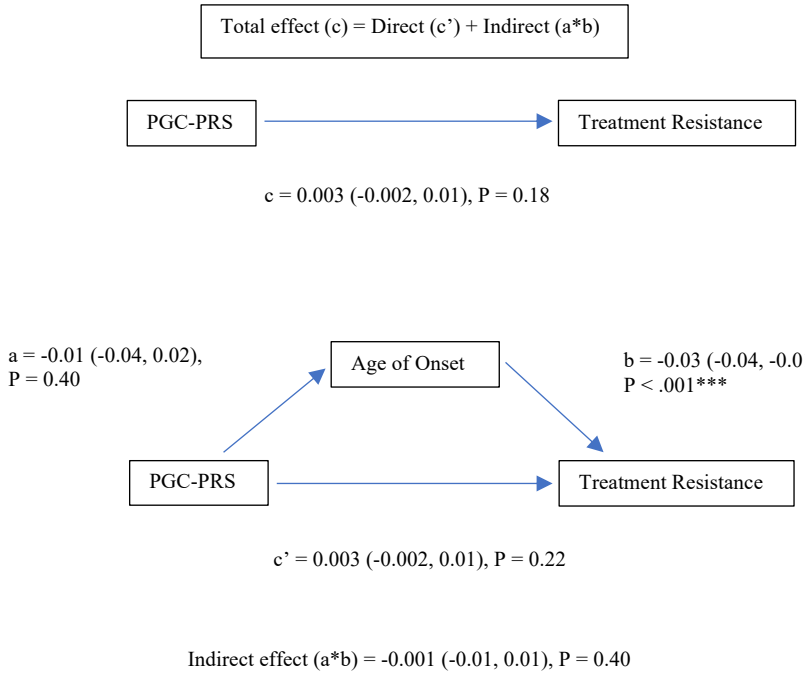


Figure 38. Output diagram showing the direct effect between the polygenic risk score for non-treatment resistance schizophrenia and treatment resistance, and the indirect effect with age of psychosis onset as a mediator.

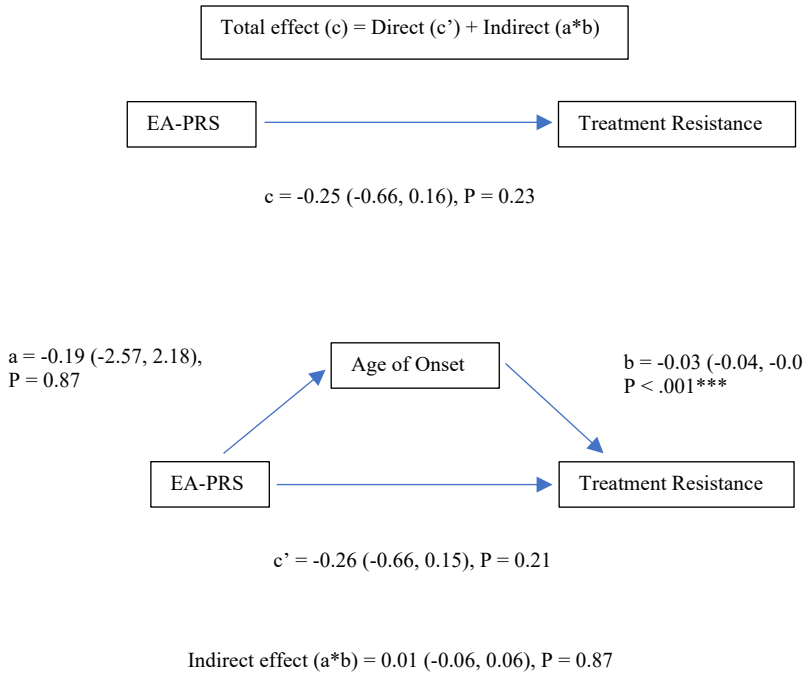


Figure 39. Output diagram showing the direct effect between the polygenic risk score for educational attainment and treatment resistance, and the indirect effect with age of psychosis onset as a mediator.

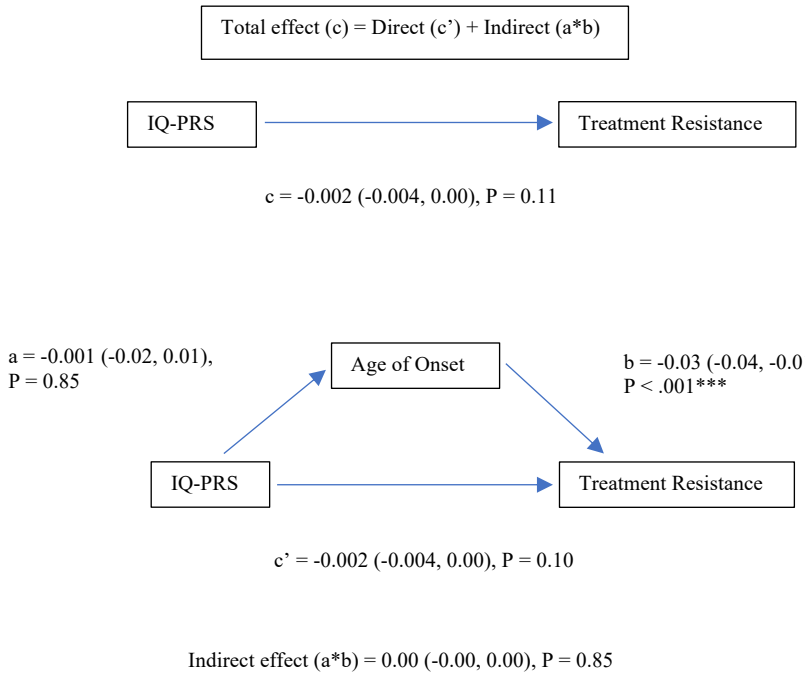


Figure 40. Output diagram showing the direct effect between the polygenic risk score for IQ and treatment resistance, and the indirect effect with age of psychosis onset as a mediator.

END OF THESIS