Pharmacogenomic scores in psychiatry: Systematic review of existing evidence

Nigussie T. Sharew¹, Scott R. Clark¹, K. Oliver Schubert¹,²,³, Azmeraw T. Amare¹*

Affiliations

¹Discipline of Psychiatry, School of Medicine, University of Adelaide, Adelaide, SA, Australia

²Northern Adelaide Local Health Network, Mental Health Services, Australia

³Headspace Adelaide Early Psychosis, Sonder, Adelaide, Australia

*Correspondence to: Azmeraw T. Amare,

Discipline of Psychiatry, Adelaide Medical School, The University of Adelaide, Adelaide, SA 5000, Australia

Tel: +61 8 83137438

E-Mail: azmeraw.amare@adelaide.edu.au

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

In the past two decades, significant progress has been made in the development of polygenic scores (PGSs). One specific application of PGSs is the development and potential use of pharmacogenomic- scores (PGx-scores) to identify patients who can benefit from a specific medication or are likely to experience side effects. This systematic review comprehensively evaluates published PGx-score studies in psychiatry and provides insights into their potential clinical use and avenues for future development.

A systematic literature search was conducted across PubMed, EMBASE, and Web of Science databases until 22 August 2023. This review included fifty-three primary studies, of which the majority (69.8%) were conducted using samples of European ancestry. We found that over 90% of PGx-scores in psychiatry have been developed based on psychiatric and medical diagnoses or trait variants, rather than pharmacogenomic variants. Among these PGx-scores, the polygenic score for schizophrenia (PGS\textsubscript{SCZ}) has been most extensively studied in relation to its impact on treatment outcomes (32 publications). Twenty (62.5%) of these studies suggest that individuals with higher PGS\textsubscript{SCZ} have negative outcomes from psychotropic treatment: poorer treatment response, higher rates of treatment resistance, more antipsychotic-induced side effects, or more psychiatric hospitalizations, while the remaining studies didn’t find significant associations. Although PGx-scores alone accounted for at best 5.6% of the variance in treatment outcomes (in schizophrenia treatment resistance), together with clinical variables they explained up to 13.7% (in bipolar lithium response), suggesting that clinical translation might be achieved by including PGx-scores in multivariable models.

In conclusion, our literature review found that there are still very few studies developing PGx-scores using pharmacogenomic variants. Research with larger and diverse populations is required to develop clinically relevant PGx-scores, using biology-informed and multi-phenotypic polygenic scoring approaches, as well as by integrating clinical variables with these scores to facilitate their translation to psychiatric practice.

Keywords: Polygenic score, Pharmacogenomics, Psychiatric disorders, Psychotropic drugs, treatment outcomes
Introduction

Psychiatric disorders are significant contributors to the global disease burden and represent a major public health concern, highlighting the urgent need for effective prevention and treatment strategies. The 2022 World Health Organization (WHO) report estimates that nearly a billion people suffer from psychiatric disorders, with an associated economic loss of $2 trillion per year and this figure is expected to rise to $6 trillion by 2030.

Pharmacological treatments including antidepressants, antipsychotics, mood stabilizers, and anxiolytics are commonly prescribed for people suffering from psychiatric disorders. However, the effectiveness of these medications varies between individuals, with some responding well while others do not show notable improvement or experience adverse effects. For example, among patients with major depressive disorder (MDD), 30-40% fail to respond to the first-line pharmacological treatment options of selective serotonin reuptake inhibitors (SSRIs), and 10-45% exhibit moderate to severe treatment-related side effects.

Similarly, only 30% of patients with bipolar disorder (BD) show a full clinical response to first-line lithium monotherapy, and up to 25% of patients with first-episode schizophrenia (SCZ) are treatment-resistant to first-line antipsychotics. This variability in pharmacological treatment outcomes can be attributed to the complex interplay of genetic and environmental factors, including patients’ clinical characteristics (e.g., severity, number, and duration of illness episodes), as well as sociodemographic variables. For example, in individuals with MDD, genetic factors account for 42-52% of the observed differences in antidepressant treatment response, while environmental factors contribute to the remainder.

To date, studies employing both candidate gene investigations (pharmacogenetics) and genome-wide (pharmacogenomic) approaches, have successfully pinpointed genetic
variations associated with treatment outcomes in psychiatry, including response, remission, resistance and adverse drug reactions. For instance, the pharmacogenetic approach has uncovered genetic polymorphisms within genes encoding drug-metabolizing enzymes including those involved in the metabolism of various psychotropic drugs (e.g., CYP2D6 and CYP2C19) as well as drug transporters (e.g., 5-HTTLPR), establishing their association with patients' responses to medications. This evidence is now incorporated into commercially available pharmacogenetic testing panels, aiding drug selection and dose adjustments and ultimately aiming at improving medication efficacy and tolerability.

Similarly, the pharmacogenomics approach has revealed a number of genetic polymorphisms located within or near pharmacologically relevant candidate genes that influence individuals' reaction to psychiatric medications. For instance, Hou et al identified four linked genetic variants on chromosome 21 associated with lithium response in a Genome-wide Association Study (GWAS). It has been challenging, however, to translate these pharmacogenomic findings into clinical practice, mainly due to the small effect size of individual genetic variants on treatment outcomes, along with a limited understanding of gene function.

In an effort to improve effect estimates and make pharmacogenomic findings more clinically relevant, researchers have recently adopted polygenic score methods combine the effect of multiple genetic variants across the genome and have developed pharmacogenomic scores (PGx-scores). In this systematic review, we provide a detailed account of the research undertaken to date, and of the performance, shortfalls, and future recommendations for the development of PGx-scores for the personalisation of psychiatric care.

**Methods**

This systematic review adhered to the PRISMA updated guidelines 2020 and was registered with the International Prospective Register of Systematic Reviews (PROSPERO).
on February 9, 2023 (ID = CRD42023395404). The review protocol was prepared before commencement to ensure a transparent and standardized methodology.

**Search strategy, inclusion, and exclusion criteria**

The literature search was performed across three databases including PubMed, EMBASE, and Web of Science databases from January 1st, 2005 to 22nd August 2023, by using search string: (("Polygenic score*" OR "Polygenic risk score*" OR "Risk profile score*" OR "Genetic risk score*" OR "Gene score*" OR "Genetic score*" OR polygenic* OR "Pharmacogenomic variants" OR "Pharmacogenomic testing" OR Pharmaco-omic* OR pharmacogeno* OR "Pharmacogenetics") AND ("Antipsychotic agents" OR antipsyc* OR "Antidepressive agents" Antidepress* OR "Anti-anxiety agents" OR Anti-anxiet* OR Valproic acid OR Valproate OR Divalproate OR Divalproex OR Carbamazepine OR Oxcarbazepine OR Risperidone OR Gabapentin OR Lamotrigine OR Licarbazepine OR Pregabalin OR Tiagabine OR Zonisamide OR Lithium))

Our search strategy included all original studies that developed PGx-score for drug-related phenotypes such as, drug dosage, therapeutic drug response, resistance, drug-induced side-effects, relapse or hospitalisation in psychiatry. We included studies that reported weighted PGx-score for the drug-related phenotypes mentioned above, while excluding publications in languages other than English, conference abstracts, case reports, editorials, notes, and systematic reviews. NTS screened the studies for inclusion under supervision of ATA. In the final step, all studies were imported into Endnote version 20, a reference manager software. Duplicate entries were removed, and the selection of studies was carried out based on the predetermined inclusion and exclusion criteria. Supplementary file 1 provides details of the systematic search strategies and results in each database.
Data extraction and synthesis

NTS extracted data using a customised data extraction excel sheet format, under supervision of ATA. This excel sheet included information on the authors' characteristics, details of the drug outcomes, characteristics of the study cohort (such as base, target, and validation cohorts), number of variants included in the polygenic score (PGS), polygenic scoring methods, and association effect estimates. The "target cohorts" describe the cohorts where the PGS was developed and tested, while "discovery cohorts" refers to the cohorts utilized to create GWAS summary statistics. "Validation cohorts" are independent cohorts where the PGSs were validated. "Variance explained" measures the proportion of phenotype variance the PGS can account for in a predictive model assuming linear effects. Coefficient of effect estimates, standard error, and sample size were used to calculate odd ratios if not reported in the studies. The results were organized thematically based on the psychiatric disorders that were studied, as well as the specific phenotypes investigated, including treatment response, treatment resistance, and drug-induced side effects.

Quality assessment

The quality of included studies was assessed using a quality assessment form adapted from previously validated and published sources. The assessment criteria covered various aspects of the study design, such as the rationale and methods of PGS, power calculation, inclusion and exclusion criteria, basic characteristics of the study population, availability of validation cohort, type of analysis, correction for multiple testing, and consideration of confounders in the analysis. The quality assessment was conducted by NTS under supervision of ATA.
Results

Our initial search identified a total of 4,889 studies that were potentially relevant to the research topic. After removing 1,586 duplicated publications, 3,303 articles remained for the title and abstract screening. Subsequently, 3,175 studies were excluded during the initial title and abstract screening phase, leaving 127 articles for full text review. Finally, 53 studies met the predetermined inclusion criteria and were included in the final synthesis. Figure 1 presents the flowchart of the step-by-step process of study selection with reasons for exclusion.

Figure 1. PRISMA flow diagram showing the steps of screening studies included in this systematic review. Abbreviations: PGS = Polygenic score; PGx = Pharmacogenomics
Quality assessment

Nearly three-quarters (39/53) of studies described the rationale for the selected polygenic scoring methods, while about 20% (10/53) of studies performed a power calculation. All studies reported the inclusion and exclusion criteria for participants' selection. Only fourteen studies used external cohorts to validate their findings. Correction for multiple testing was performed in 83.2% (44/53) of studies. Detailed results of the quality assessment are provided in a supplementary file 1.

Most studies 37 (69.8%) were conducted on samples comprising individuals of European ancestry. Eleven studies (20.8%) included participants from other ancestries, such as African, African American and/or East Asian. Three studies targeted only Latin American participants and another two studies were conducted specifically on samples of East Asian ancestry. However, there was no study solely centered on samples of African ancestry. A combined analysis of both the target and discovery samples showed that 14,893,321 (90%) of participants had European descent, with an increased trend over the years 2013-2023, both in the target (Figure 2A) and discovery cohorts (Figure 2B).
Figure 2. The ancestry characteristics of samples in the A) target and B) discovery cohort for studies from 2013 – 2023.

The sample sizes across the studies varied widely, ranging from 44 participants \(^{22}\) to 12,863 participants \(^{30}\) with a median sample size of 863 in the target cohorts. Three major psychiatric conditions namely SCZ, depression and BD were the focus of included studies. In the case of SCZ, nearly 80% (21/27) of studies investigated the association between PGS and response to second-generation antipsychotics (clozapine, risperidone, lurasidone, olanzapine, aripiprazole, quetiapine, ziprasidone, and perphenazine). About half of SCZ studies (13/27)
exclusively analysed clozapine treatment outcome. Nearly three-quarters of the included studies involving patients MDD (14/19) considered the relationship between PGS and SSRIs such as citalopram or escitalopram. Six out of seven included studies developed PGx-scores and examined their associations with lithium treatment response in patients with BD.

The association of pharmacogenomic scores with treatment outcomes

Involving patients with psychiatric disorders, researchers have developed several PGx-scores and investigated their associations with key treatment outcomes: treatment response, treatment resistance, treatment-related side effects, and hospitalization rates. Table 1 provides a summary of the findings extracted from each of the articles included in this systematic review (Table 1).
Table 1. Summary of findings on the association between PGx-score and treatment outcomes.

<table>
<thead>
<tr>
<th>Author</th>
<th>Target cohort characteristics, pharmacogenomic- scores and treatment outcomes</th>
<th>Associations of polygenic scores with treatment outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cohort</td>
<td>N</td>
</tr>
<tr>
<td>Guo et al.,</td>
<td>CAPOC and CAPEC</td>
<td>3686</td>
</tr>
<tr>
<td>(2023)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kappel et al.,</td>
<td>CLOZUK</td>
<td>3133</td>
</tr>
<tr>
<td>(2023)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santoro et al.,</td>
<td>CAISM</td>
<td>60</td>
</tr>
<tr>
<td>(2018)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Talarico et al.,</td>
<td>PROESQ</td>
<td>174</td>
</tr>
<tr>
<td>(2022)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al.,</td>
<td>ZHH-FE</td>
<td>77</td>
</tr>
<tr>
<td>(2019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gasse et al.,</td>
<td>DCRS</td>
<td>593</td>
</tr>
<tr>
<td>(2019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lin et al.,</td>
<td>GROUP and CLOZIN</td>
<td>2505</td>
</tr>
<tr>
<td>(2023)</td>
<td></td>
<td>687</td>
</tr>
<tr>
<td>Werner et al.,</td>
<td>TOP</td>
<td>321</td>
</tr>
<tr>
<td>(2020)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>N</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>Pardinas et al., (2022)</td>
<td>Cardiff COGS and STRATA-G</td>
<td>817</td>
</tr>
<tr>
<td>Pain et al., (2022)</td>
<td>STAR*D, GSRD, GENDEP, DAST, PGRN-AMPS, GENPOD, PFZ, Mayo, GSK, GODS, Miaoli, Taipei, Japan</td>
<td>5117</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facal et al., (2022)</td>
<td>GEHRS</td>
<td>1241</td>
</tr>
<tr>
<td>Ward et al., (2018)</td>
<td>PGRN-AMPS and GENDEP</td>
<td>1065</td>
</tr>
<tr>
<td>Amare et al., (2021)</td>
<td>ConLi`Gen</td>
<td>2586</td>
</tr>
<tr>
<td>Gendep Investigators et al (2013)</td>
<td>GENDEP, STAR*D and MARS</td>
<td>2256</td>
</tr>
<tr>
<td>Schubert et al., (2021)</td>
<td>ConLi`Gen</td>
<td>2283</td>
</tr>
<tr>
<td>Mayen-Lobo et al., (2021)</td>
<td>NINN</td>
<td>44</td>
</tr>
<tr>
<td>Coombes et al., (2021)</td>
<td>ConLi`Gen</td>
<td>2510</td>
</tr>
<tr>
<td>Blackman et al., (2022)</td>
<td>Healthy volunteers</td>
<td>71</td>
</tr>
</tbody>
</table>

All rights reserved. No reuse allowed without permission.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Study Design</th>
<th>N</th>
<th>Primary Outcome Measure</th>
<th>Treatment/Phenotype</th>
<th>Comparison</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yoshida et al., (2023)</td>
<td>CATIE and CAMH</td>
<td>151 and 138</td>
<td>SCZ and T1D</td>
<td>Second generation antipsychotics</td>
<td>AIWG</td>
<td>Higher PGS for T1D compared with lower: 4.67 [1.66, 13.12]; Higher PGS for SCZ compared with lower: 2.04 [1.14, 3.66]</td>
</tr>
<tr>
<td>O’Connell et al., (2023)</td>
<td>TDM</td>
<td>1733</td>
<td>SCZ, TRS and BMI</td>
<td>Clozapine</td>
<td>TRS</td>
<td>NR</td>
</tr>
<tr>
<td>Millischer et al., (2022)</td>
<td>SWEBIC</td>
<td>2357</td>
<td>BMI and BUN</td>
<td>Lithium</td>
<td>CL&lt;sub&gt;Li&lt;/sub&gt;</td>
<td>Positive association between PGS&lt;sub&gt;BMI&lt;/sub&gt; and CL&lt;sub&gt;Li&lt;/sub&gt;; and negative association between PGS for BUN with CL&lt;sub&gt;Li&lt;/sub&gt;</td>
</tr>
<tr>
<td>Amare et al., (2019)</td>
<td>STAR*D and ISPC</td>
<td>865</td>
<td>CAD and Obesity</td>
<td>SSRIs</td>
<td>Response</td>
<td>Fourth PGS for CAD quartile: first in ISPC: 0.71, [0.52-0.96]; Fourth PGS for obesity quartile: first in ISPC: 0.53, [0.32-0.88]</td>
</tr>
<tr>
<td>Men et al., (2023)</td>
<td>STOP-PD II</td>
<td>205</td>
<td>Antidepressant response and Alzheimer’s disease</td>
<td>Sertraline and Olanzapine</td>
<td>Remission and relapse</td>
<td>PGS for antidepressants symptom improvement with remission status: 1.95 [1.20, 3.17]; PGS for Alzheimer’s disease with relapse: 0.38 [0.18, 0.80]</td>
</tr>
<tr>
<td>Campos et al., (2022)</td>
<td>AGDS</td>
<td>12,863</td>
<td>Chronic pain and MDD</td>
<td>SSRIs</td>
<td>Response</td>
<td>Higher PGS pain group vs lower: 0.95 [0.92, 0.98]</td>
</tr>
<tr>
<td>Marshe et al., (2021)</td>
<td>IRL-GRAY</td>
<td>355</td>
<td>MDD, Alzheimer’s disease and cardioembolic disease</td>
<td>Venlafaxine</td>
<td>Late-life treatment response</td>
<td>Higher PGS for cardioembolic stroke with non-remission vs lower: 0.63 [0.48, 0.83]</td>
</tr>
<tr>
<td>Study</td>
<td>Design/Drug</td>
<td>Sample Size</td>
<td>Response Measure</td>
<td>Effect Size (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hommers et al., (2021)</td>
<td>RPUHW</td>
<td>804</td>
<td>QT intervals</td>
<td>Higher PGS for QT interval vs lower: 1.19 [1.02, 1.44]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amare et al., (2023)</td>
<td>ConLi’Gen</td>
<td>2367</td>
<td>Lithium responsiveness</td>
<td>Tenth Li’R&lt;sub&gt;PGS&lt;/sub&gt; decile vs first Li’R&lt;sub&gt;PGS&lt;/sub&gt; decile for favourable response: 3.47 [2.22-5.47]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meijs et al., (2022)</td>
<td>ZNA and iSPOT-D</td>
<td>1123</td>
<td>Antidepressants response</td>
<td>Higher PGS for antidepressants response vs lower: OR (1.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muntane et al., (2023)</td>
<td>PAFIP</td>
<td>381</td>
<td>SCZ and BMI</td>
<td>Higher PGS for BMI and SCZ vs lower: OR (1.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morgenroth et al., (2023)</td>
<td>CLOZIN</td>
<td>102</td>
<td>SCZ and OCD</td>
<td>NSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segura et al., (2022)</td>
<td>CIBERSAM</td>
<td>231</td>
<td>SCZ, BD, depression and T2D</td>
<td>~1.20 – 4.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lu et al., (2022)</td>
<td>CAPOC</td>
<td>2040</td>
<td>Myocardial Infarction</td>
<td>Higher PGS for QTc prolongation was associated with increased antipsychotic induced QTc interval prolongation</td>
<td>~0.01</td>
<td></td>
</tr>
<tr>
<td>Kowalec et al., (2021)</td>
<td>SNPD</td>
<td>4936</td>
<td>SCZ, MDD and BD</td>
<td>NSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lacaze et al., (2020)</td>
<td>Registered patients</td>
<td>109</td>
<td>Myocarditis</td>
<td>NSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maciukiewicz et al., (2019)</td>
<td>CAMH</td>
<td>201</td>
<td>BMI and obesity</td>
<td>NSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cearns et al., (2022)</td>
<td>ConLi+Gen</td>
<td>1034</td>
<td>SCZ and MDD</td>
<td>Lithium</td>
<td>Response</td>
<td>Not specified</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>----------</td>
<td>--------------</td>
</tr>
<tr>
<td>Li et al., (2018)</td>
<td>Registered patients</td>
<td>302</td>
<td>SCZ</td>
<td>Lurasidone</td>
<td>Response</td>
<td>PGS&lt;sub&gt;SCZ&lt;/sub&gt; was associated with the exacerbation of positive symptoms of SCZ.</td>
</tr>
<tr>
<td>Wimberley et al., (2017)</td>
<td>DPR</td>
<td>862</td>
<td>SCZ</td>
<td>Clozapine</td>
<td>TRS</td>
<td>NSA</td>
</tr>
<tr>
<td>Martin &amp; Mowry, (2016)</td>
<td>MGS</td>
<td>612</td>
<td>SCZ</td>
<td>Clozapine</td>
<td>TRS</td>
<td>NSA</td>
</tr>
<tr>
<td>Heitige et al., (2016)</td>
<td>CAMH</td>
<td>83</td>
<td>SCZ</td>
<td>Second generation antipsychotics</td>
<td>Response</td>
<td>NSA</td>
</tr>
<tr>
<td>Nøhr et al., (2022)</td>
<td>Seven clinical trials</td>
<td>1364</td>
<td>MDD, BD, SCZ</td>
<td>Vortioxetine</td>
<td>Response</td>
<td>NSA</td>
</tr>
<tr>
<td>Fanelli et al., (2022)</td>
<td>Brescia, GSRD, Münster, STAR*D, Tartu</td>
<td>3637</td>
<td>SCZ, BD, MDD and neuroticism</td>
<td>SSRIs</td>
<td>Non-response</td>
<td>Higher PGS&lt;sub&gt;MDD&lt;/sub&gt; vs lower: 1.10 [1.02–1.19] (Nominally)</td>
</tr>
<tr>
<td>Fanelli et al., (2022)</td>
<td>Same cohorts</td>
<td>3184</td>
<td>Non-remission</td>
<td>Higher PGS&lt;sub&gt;MDD&lt;/sub&gt; vs lower: 1.14 [1.04–1.24] (Nominally)</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Taylor et al., (2021)</td>
<td>Bethlem Royal Hospital</td>
<td>240</td>
<td>MDD</td>
<td>Mix of antidepressants</td>
<td>Treatment resistance</td>
<td>NSA</td>
</tr>
<tr>
<td>Fanelli et al., (2021)</td>
<td>GSRD</td>
<td>1148</td>
<td>SCZ, BD, MDD and neuroticism</td>
<td>SSRIs</td>
<td>Treatment non-response</td>
<td>Higher PGS&lt;sub&gt;SCZ&lt;/sub&gt; vs the lower: 2.23: [1.21–4.10]</td>
</tr>
<tr>
<td>Wigmore et al., (2020)</td>
<td>GS: SFHS</td>
<td>3452</td>
<td>SCZ, MDD, and BD</td>
<td>Mixed antidepressants</td>
<td>Treatment resistance</td>
<td>PGS&lt;sub&gt;MDD&lt;/sub&gt; = 1.01; PGS&lt;sub&gt;SCZ&lt;/sub&gt; = 1.01; PGS&lt;sub&gt;BD&lt;/sub&gt; = 1.01 (Nominally)</td>
</tr>
<tr>
<td>Li et al., (2020)</td>
<td>SUSTAIN-2 and TRANSFORM-3</td>
<td>527</td>
<td>SCZ, MDD and BD</td>
<td>Esketamine</td>
<td>Response</td>
<td>NSA</td>
</tr>
<tr>
<td>Garcia-Gonzalez et al., (2017)</td>
<td>GENDEP, STAR*D, GENPOD, GODS, GSK, Pfizer, Muenster</td>
<td>3746</td>
<td>SCZ and MDD</td>
<td>SSRIs</td>
<td>Response</td>
<td>NSA</td>
</tr>
<tr>
<td>Tansey et al., (2014)</td>
<td>NEWMEDS &amp; STAR*D</td>
<td>2897</td>
<td>BD</td>
<td>SSRIs</td>
<td>Response</td>
<td>NSA</td>
</tr>
</tbody>
</table>
Abbreviations: PGx-score = Pharmacogenomic polygenic score; PGS = Polygenic score; SCZ = Schizophrenia; MDD = Major Depressive Disorders; MDE = Major Depressive Episode; BD = Bipolar Disorders; ADHD = Attention Deficit Hyperactivity Disorders; OCD = Obsessive-Compulsive Disorders; PGS_{SCZ} = PGS for SCZ; PGS_{MDD} = PGS for MDD; PGS_{BD} = PGS for BD; CAPOC = Chinese Antipsychotics Pharmacogenomics Consortium; CAPEC = Chinese Antipsychotics Pharmacogenetics Consortium; CLOZUK = Genome-wide genotype information for SCZ cases from the UK; CAISM = Centro de Atención Integral a Saúde Mental; PROESQ = Schizophrenia Program at the Universidade Federal de São Paulo; ZHH-FE = Zucker Hillside Hospital First Episode schizophrenia trial; DCRS = Danish Civil Registration System; CLOZIN = Clozapine International; GROUPS = Genetic Risk and Outcome of Psychosis; TOP = Thematically Organized Psychosis; Cardiff COGS = Cardiff Cognition in Schizophrenia; STRATA-G = Genetics Workstream of the Schizophrenia Treatment Resistance and Therapeutic Advances; GSRD = Group for the Study of Resistant Depression; GENDEP = Genome Based Therapeutic Drugs for Depression; DAST = Depression and Sequence of Treatment; PGRN-AMPS = Pharmacogenomics Research Network Antidepressant Medication Pharmacogenomic Study; GENPOD = Genetics and clinical Predictors of treatment response in depression; GODS = Geneva Outpatient Depression Study; CAMH = Centre for Addiction and Mental Health; PAFIP = Cantabria program for early interventions in psychosis; STOP-PD II = Study of pharmacotherapy of psychotic depression II; GEHRS = Galician electronic health records system; ConLi^*Gen = International Consortium of Lithium Genetics; MARS = Munich Antidepressants Response Signature; CATIE = Clinical Antipsychotic Trials of Interventions Effectiveness; TDM = Therapeutic drug monitoring; NINN = National Institute of Neurology and Neurosurgery Manuel Velasco Suárez; DPR = Danish population-based registers; SWEBIC = Swedish bipolar cohort; MGS = Molecular Genetics of Schizophrenia; STAR*D = Sequenced Treatment Alternatives to Relieve Depression; AGDS = Australian Genetics of Depression Study; IRL-GREY = Incomplete Response in Late Life Depression; Getting to Remission; RPUHW = Registered patients at the University Hospital of Würzburg; IQ = Intelligent Quotient; BUN = Blood Urea Nitrogen; CL_{Li} = Total body lithium clearance; proxyDNAm = Proxy DNA methylation; PGS_{pain} = Polygenic score for pain; PGS_{cog} = Polygenic score for cognitive ability; Li^*R_{PGS} = PGS for lithium responsiveness; CRP = C-reactive protein; TRS = Treatment resistant Schizophrenia; SNPD = Swedish National Prescribed Drug Register; BMI = Body Mass Index; HDL = High-density lipoprotein; LDL = Low-density lipoprotein; TG = Triglyceride; TC = Total cholesterol; T1D = Type 1 Diabetes; T2D = Type 2 Diabetes; CAD = Coronary Artery Diseases; SSRIs = Selective Serotonin Reuptake Inhibitors; GS = Glaxo Smith Kline; PFZ = Pfizer; ISP = International SSRI Pharmacogenomics Consortium; ZNA = Ziekenhuis Netwerk Antwerpen; iSOT-D = International Study to Predict Optimized Treatment in Depression; GS-SFHS = Generation Scotland: the Scottish Family Health Study; SUSTAIN-2 = Long-term Safety and Efficacy Study of Intranasal Esketamine in Treatment-resistant Depression; TRANSFORM = Safety and Tolerability of Intranasal Esketamine Plus Oral Antidepressant in Elderly Participants with Treatment-resistant Depression; NEWMEDS = New Medications in Depression and Schizophrenia; HR = Hazard Ratio; SD = Standard Deviation; NSA = No significant association; NR = Not reported; NA = Not applicable; CI = Confidence Interval
Of all PGx-scores, the polygenic loading for schizophrenia (PGS\textsubscript{SCZ}) has been most extensively studied (32 publications) in relation to its influence on treatment outcomes. Several studies (20 studies) revealed that individuals with higher genetic loading for SCZ had a poorer treatment outcome to psychiatric medications \cite{31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50}, while the remaining studies didn’t find significant associations. For example, in patients with SCZ, a negative correlation ($r = -0.05$ [95\%CI: $-0.09$ – $-0.01$]) was found between PGS\textsubscript{SCZ} and response to second generation antipsychotics (olanzapine, aripiprazole, risperidone, quetiapine, haloperidol, ziprasidone, perphenazine) following six weeks treatment \cite{44}. Kappel et al.\cite{2023} observed a positive correlation ($\beta=12.21$; 95\%CI: $4.81 – 19.62$) between PGS\textsubscript{SCZ} and high clozapine dosing (>600 mg/day) suggesting that individuals with a higher PGS\textsubscript{SCZ} may require increased doses of clozapine to achieve effective treatment response \cite{33}. In patients treated with risperidone, those who had a higher PGS\textsubscript{SCZ} reported more depressive symptoms \cite{45}, and worsened positive and negative psychotic symptoms \cite{46}. A higher PGS\textsubscript{SCZ} was associated with a poor response to olanzapine or risperidone OR=1.43 [95\%CI:1.19 – 1.67] \cite{43} and an increase of one standard deviation in PGS\textsubscript{SCZ} was associated with an approximately 11\% increase in the risk of developing treatment-resistant schizophrenia (TRS) (OR = 1.11[95\%CI: 1.00 – 1.24]) \cite{44}. Patients with a higher polygenic load for SCZ were 1.22 times [95\%CI: 1.05 – 1.41; R\textsuperscript{2} = 2.03\%] more likely to be resistant to clozapine \cite{36}, had 1.50 times [95\%CI: 1.13 – 1.96; R\textsuperscript{2} = 1.70\%] higher odds to experience resistance to other antipsychotics \cite{42} and were more likely to develop antipsychotic-induced weight gain (AIWG) \cite{48, 49}.

In decile-based comparisons, patients in the top three PGS\textsubscript{SCZ} deciles had a 2.42 [95\%CI: 1.35 – 3.49; R\textsuperscript{2} ~2.00\%] times higher odds of poor response to various antipsychotic medications (olanzapine, risperidone, quetiapine, and clozapine) \cite{35} and the odds of treatment resistance for those in the 8\textsuperscript{th} PGS\textsubscript{SCZ} decile was 6.50 times [95\%CI: 1.47 – 28.80] higher than for
patients in the 1st decile. Patients with a higher PGS\textsubscript{SCZ} had 1.48 times [95%CI: 1.10 - 1.97; \( R^2 = 2.70\% \)] higher odds of psychiatric hospitalizations and were hospitalized longer.

Interestingly, in a study by Okhuijsen-Pfeifer et al (2022), patients treated with clozapine who were in the highest PGS\textsubscript{SCZ} tertile group were 1.94-fold more likely to experience low (i.e., more favourable) symptom severity [95%CI: 1.33 – 2.81; \( R^2 = 1.85\% \)], compared to those in the lowest PGS\textsubscript{SCZ} tertile group.

Similar to the above findings, studies involving patients with MDD and BD have also unveiled the association of a higher PGS\textsubscript{SCZ} with poorer response to either antidepressants in MDD (OR = 0.97 [95%CI: 0.96 - 0.98; \( R^2 = 0.01\% \)] or lithium treatment in BD (OR = 0.29 [95%CI: 0.12 – 0.70; \( R^2 = 0.80\% \)]

Aside from the PGS\textsubscript{SCZ}, several other polygenic scores have also been investigated for their potential to predict treatment outcomes in patients with psychiatric disorders. These include polygenic scores for cognitive function, cardiometabolic traits, MDD, BD, ADHD, autism spectrum disorders (ASD), anxiety, alcohol use disorders, personality traits, and educational attainment. In patients with SCZ, a higher genetic loading for general cognitive ability was associated with better cognitive function following antipsychotic treatment. SCZ patients carrying a greater genetic load for higher body mass index (BMI) were at a higher risk of being resistant to clozapine treatment. Moreover, the polygenic loading for BD (PGS\textsubscript{BD}) was found to be significantly associated with clozapine metabolic ratio, a measure of how clozapine is metabolized within the body, which may impact treatment response or adverse effects. In patients with first episode psychosis, higher genetic loadings for HDL, LDL, and total cholesterol predicted antipsychotic-induced metabolic disturbance.

Studies of PGx-scores for patients with MDD also explored the impact of the polygenic scores for personality traits and physical illnesses on treatment outcomes. In a study by the
Genome-Based Therapeutic Drugs for Depression (GENDEP) investigators, the polygenic loading for MDD (PGS<sub>MDD</sub>) was significantly associated with response and remission to SSRIs and TCAs treatment in patients with MDD, although the direction of association was not reported. A study that assessed the relationship between PGS for various personality traits and response to SSRIs (citalopram, escitalopram, fluvoxamine) found that a higher genetic loading for openness personality trait was associated with a better SSRIs treatment response after 8-week of treatment OR=1.58 [95%CI, 1.10-2.90] while the PGS for neuroticism was negatively associated with SSRIs treatment response. Genetic loading for cardiometabolic disease has also shown associations with response to antidepressant treatment: Marshe et al (2021) used a PGS for cardioembolic stroke to predict response to venlafaxine, an antidepressant of the serotonin-norepinephrine reuptake inhibitors (SNRI) class, after 12-week treatment. They found that a one standard deviation increase in PGS for cardioembolic stroke was associated with a decreased probability of remission (OR = 0.63[95%CI:0.48 to 0.83]) or worsened disease symptoms (Montgomery-Asberg Depression Rating Scale (MADRS), β = -5.51[95%CI: -9.45 to -1.57]). In a different study, individuals with the highest PGSs for coronary artery disease (4th quartile) had 0.53 times ([95%CI, 0.35 – 0.81]) less likelihood of experiencing favourable response to SSRIs (citalopram, escitalopram, fluvoxamine) compared to those in the 1st quartile. Similarly, those with higher genetic loading for obesity (4th quartile) had a 0.53 times ([95%CI, 0.32 – 0.88]) lower likelihood of achieving a positive response to SSRIs treatment. In individuals with MDD (n=5218) treated with SSRIs (citalopram, escitalopram) or a TCA (nortriptyline), the PGS for educational attainment was positively associated with SSRI response. In a cohort of patients with psychotic depression treated with sertraline and olanzapine for 36 weeks, those who had a higher polygenic loading for Alzheimer’s disease had a decreased likelihood of relapse (OR = 0.38; [95%CI: 0.18 – 0.80]) during the study period. Higher PGS for chronic
pain was negatively associated with treatment response to SSRIs, tetracyclic antidepressants (mirtazapine), and SNRIs (desvenlafaxine) (OR = 0.95 [95% CI: 0.92 – 0.98]) \(^{30}\), while a higher PGS for C-reactive protein was associated with a better response to escitalopram (OR = 2.92 [95% CI: 1.30 – 6.49]), but worse response to nortriptyline \(^{38}\).

In patients with BD, those with a low PGS\textsubscript{MDD} (first decile) were 1.54 times [95% CI: 1.18 – 2.01; R\textsuperscript{2} = 0.91%] more likely to respond favourably to lithium than those who had high depression genetic loading (10\textsuperscript{th} decile) \(^{59}\). Similarily, a higher PGS for ADHD was associated with unfavourable lithium response (OR = 0.86 [95% CI: 0.77 - 0.95], R\textsuperscript{2} = 0.18) \(^{60}\). Further studies using the same dataset have shown that a combined analysis of the PGSs of multiple phenotypes and PGS with patients’ clinical data can improve the predictive capacity of polygenic models. For example, a meta-analysis of the association results of the PGS\textsubscript{SCZ} and PGS\textsubscript{MDD} provided improved response prediction compared to single disorder PGS \(^{41}\). By applying machine learning methods, the PGS\textsubscript{SCZ} and PGS\textsubscript{MDD} were combined with clinical data which resulted in an explained variance of 13.7% in lithium treatment response \(^{40}\). In a recent study, lithium clearance, an essential parameter for maintaining therapeutic levels of lithium and adjusting dosage, was positively associated with the PGSs for BMI and estimated glomerular filtration rate (eGFR), while it was negatively associated with the PGSs blood urea nitrogen (BUN) \(^{61}\).

In contrast to the above studies in which PGx-scores were developed based on diseases or related phenotype variants, a few recent studies used pharmacogenomic variants to calculate PGx-scores, directly indexing treatment outcome phenotypes. For instance, in cohort of patients with TRS, a PGx-score for clozapine resistance predicted 4.96% of the rate of TRS variation \(^{31}\). In patients with psychotic depression treated with sertraline and olanzapine, those with a higher genetic loading for antidepressant remission and response had 1.95 times...
[95%CI: 1.20 – 3.17] higher odds of reaching remission after 36 weeks. In similar context, PGS for response to SSRI (escitalopram, sertraline, venlafaxine) predicted antidepressant treatment response in patients with MDD. A study by Guo et al. (2018) utilized variants ranked by their strength of association with ketamine response, a glutamate-modulating antidepressant used in patients with Treatment-Resistant Depression (TRD), to predict scopolamine treatment response in patients with either MDD or BD who had a current major depressive episode. Findings indicated that patients with higher genetic loadings for ketamine response had better responses to scopolamine, an emerging antidepressant with effects on acetylcholine (Ach) neurotransmission. A polygenic score developed for lithium treatment response (Li+RPGS) within ConLi+Gen was evaluated in a hold-out subsample and a smaller independent replication cohort. This analysis revealed that individuals in the highest Li+RPGS decile were 3.47 times [95%CI: 2.22 – 5.47, R² = 2.60] more responsive to lithium compared to those in the lowest PGS decile, and a linear relationship was observed across the various deciles. Figure 3 summarized the relationship between different PGx-scores of different traits and pharmacotherapeutic outcomes.
Figure 3. Figure showing the relationship between different PGx-scores of different traits and pharmacotherapeutic outcomes in psychiatry.

Legend: Green line represents the positive associations of PGx-scores with treatment outcomes; Gray line indicates negative associations between PGx-scores and treatment outcomes. A wider (thick) line lines represent a stronger association.
Discussion

Pharmacogenomic scores (PGx-scores) are emerging as novel tools for predicting treatment outcomes in psychiatry such as response, remission, resistance, side effects, or hospitalization rates. While the bench-to-bedside translation of PGx-scores has not yet been achieved, a growing body of evidence indicates their potential clinical use for treatment personalisation. In this systematic review, we describe the landscape of 53 PGx-score studies in clinical psychiatry. These PGx-scores have been developed either from genetic variants associated with psychiatric or medical diagnoses (the majority of studies); or from pharmacogenomic variants associated with treatment outcome phenotypes (a few recent studies). Individual PGx-scores alone do not explain enough variance in clinically relevant outcomes and their combination with clinical data and/or other biological markers is required for effective translation.

First, we found that over 90% of PGx-scores have been developed based on genetic variants of psychiatric or medical diagnoses (e.g., SCZ, MDD, BD, ADHD, coronary artery disease (CAD)) or phenotypes related to diagnoses (e.g., cognitive function, personality traits, educational attainment, CRP level, BMI). Among these, the PGS\textsubscript{SCZ} has been most extensively studied and has consistently shown association with pharmacotherapeutic outcomes across drug classes including antipsychotics, antidepressants and lithium, explaining as much as 3.2% of interindividual variability in some treatment outcomes \(^{43}\). The consistent association of the PGS\textsubscript{SCZ} and treatment outcomes may be attributed to two factors. First, SCZ has a strong genetic basis with a heritability estimate of 80-85\% \(^{65}\) and it is possible that PGS\textsubscript{SCZ} captures a substantial amount of the phenotypic variance of the disorder. Previous studies have shown a direct correlation between a higher phenotypic heritability and a better predictive power of PGS \(^{66}\). Second, SCZ GWASs are well powered, including cases and controls of diverse ancestral background \(^{67, 68}\), leading to more accurate PGSs \(^{69}\). The size of GWAS discovery samples has been associated with a better accuracy and predictive power of PGs \(^{69}\). For
example, the Psychiatric Genomics Consortium (PGC in 2009) found that common genetic variants explained only 3% of the total variance in risk to SCZ, in a sample of 3322 individuals with SCZ and 3587 controls of European ancestry. In a follow up study (in 2014) with expanded sample size and diversity (36,989 cases, 113,075 controls, multiple cohorts of East Asian ancestral background), the variance explained by PGS\textsubscript{SCZ} substantially increased to around 18%.

It is important to highlight that in most of the reviewed studies, high PGS\textsubscript{SCZ} was associated with poor treatment response, more treatment resistance, more antipsychotics-induced side effects, or more psychiatric hospitalizations. A notable exception was a positive association with lower symptom burden in SCZ patient treated with clozapine. A possible explanation is that a high PGS\textsubscript{SCZ} loadings may index individuals with a higher neurodevelopmental contribution to mental disorder aetiology. Neurodevelopmental hypotheses are well established in SCZ, for instance the excessive synaptic pruning linked to complement system genotype. Psychosis prodrome and onset and TRS has been linked to reduced brain volume and connectivity. These ‘hard wired’ brain characteristics may be more difficult to influence therapeutically through first-line (e.g., non-clozapine) pharmacological strategies.

The review also identified polygenic associations between cardiometabolic disorders, personality traits, and treatment outcomes. A higher PGSs for CAD, obesity, and neurotic personality were associated with poor response to antidepressants while a positive association was found with the PGS for openness personality. This is possibly due to shared biological mechanisms, for example, a genetic overlap between major psychiatric disorders and cardiometabolic diseases, neuroticism, or openness personality traits and also associated multimorbidity across these disorders that might impact patients treatment outcomes. Personality traits have an impact on medication adherence, with
neuroticism linked to non-adherence and openness to compliance 44. These findings indicated that disease related PGS may help us to understand underlying pathology and identify drug targets, however, there may be limitations in their utility for pharmacogenomic testing due to challenges of interpretation.

Second, from our review, it is clear that there is a major research gap regarding PGx-scores developed from pharmacogenomic variants 31, 57, 62,64. The lack of these studies is associated with the limited availability of well powered GWAS summary statistics on treatment outcomes (target sample) and challenges to collect genetic and clinical data from patients of a specific diagnoses, treated with similar medications (discovery sample). Currently, large scale GWASs leverage biobank datasets, where there is limited phenotyping on medication, missing standardised data on treatment outcomes.

Although the current cohort sizes for PGx-score development are much smaller than those of large scale diagnosis-based GWASs, promising initiatives are underway to achieve deeper phenotyping for medications such as lithium 85, clozapine 36, 86, and antidepressants 37. For instance, ConLi+Gen cohort, which aimed to study the genetics of lithium treatment response in individuals with BD, currently has a sample size of 2367 patients of European ancestry and 220 patients of Asian ancestry with current effort underway for a larger more diverse cohort and more detailed phenotyping 85. By expanding current efforts, there may be opportunities to develop pharmacogenomic PGx-scores with improved accuracy for clinical use.

The third finding from this review is that PGx-score alone fall short of explaining adequate variance in treatment outcomes for clinical translation. Notably, the highest reported explained variance solely attributed to PGx-score, by leveraging genetic variants of TRS and BMI, was 5.6% in resistance to clozapine. To address this shortfall, the combination of PGx-scores with clinical data could potentially enhance clinical use. For instance, a study modelled PGS_{SCZ} +
PGS\textsubscript{MDD} with patients' clinical characteristics using machine learning, was able to explain 13.7\% of the variance in lithium treatment responses \cite{40}. A further example is a multimodal model combining PGS with socio-demographic, clinical, biomarkers and structural imaging to predict rehospitalization risk showed a negative predictive value of 81.57\% compared with a PGS-only model (54.83\%) \cite{87}. Similarly, a study that modelled polygenic scores of SCZ, MDD, and BD, along with proxy DNA methylation data and clinical symptom variables showed good regression performance for prediction of response to multiple antipsychotic drugs (ROC = 0.87 [95\% CI: 0.87-0.88] \cite{34}. In patients with type 2 diabetes, combining PGS with clinical data such as smoking status, BMI, blood lipid levels, blood pressure, and the use of anti-hypertensive and lipid-lowering medications, substantially improved the accuracy in classifying individuals into low-, moderate-, and high-risk categories for cardiovascular events to 83\%, whereas accuracy was 58\% with PGSs alone (29 optimized univariable PGS) \cite{88}. It is evident from these studies that PGx-score can be clinically useful if prediction models are refined based on a combination of PGx-scores and clinical data.

**Limitations**

Some of the limitations of the present systematic review should be highlighted. First, the study participants of the studies were predominantly drawn from European populations that limits the ability to apply study’s conclusion to non-European populations and raising concerns about the generalizability of the findings to more diverse populations. Second, the inconsistent reporting of the polygenic model parameters across studies makes it challenging to compare the efficacy and obscure the true picture of PGx-score in predicting psycho-pharmacotherapeutic outcomes. Third, a significant portion of the included studies lack sufficient statistical power to draw conclusive results to the broader populations. Finally, the lack of a standard definition of pharmaceutical outcomes, differences in participants
characteristics, and the use of multiple medications across the different studies makes it difficult to compare findings and to perform meta-analysis.

Where associations between PGx-scores and treatment outcomes were established, effect size estimates (beta, odds ratios, hazard ratios) and measures of explained variance ($R^2$) varied widely in studies included in our systematic review. For instance, the $R^2$ of PGx-score models for predicting resistance to clozapine treatment with PGS$_{SCZ}$ in TRS individuals ranged from 2.03% to 5.62%. Similarly, the reported odds ratios for clozapine response ranged from 1.94 [95%CI: 1.33 – 2.81] to 6.50 [95%CI: 1.47 – 28.80]. These inconsistent findings can partly be explained by phenotypic heterogeneity, evident in diverse definitions and measurement of treatment outcomes and by differences in the sample size of these studies. As an example, the definition of TRS and TRD varies widely across studies. Achieving uniformity in phenotype characterization, and harmonizing assessments across studies would help to improve the reliability of PGx-score for treatment outcomes.

Variation in sample size can also affect the size of individual study effect estimates and their statistical significance. Studies with small target or discovery samples, have limited statistical power to detect significant associations. Choi et al have demonstrated that in a discovery cohort of 100,000 samples, 200 to 500 samples in the target cohort are requisite to achieve 80% power for predicting traits across a spectrum of heritability estimates ($h^2$:0.11 – 0.23) in polygenic models. Smaller sample sizes lead to larger sampling variance on individual marker effects and error accumulates across multiple markers such that the sample of variation on polygenic scores can be considerable. Recruiting a sufficiently large and well-characterized sample of uniformly treated individuals is a common challenge in PGx-score studies.
Future directions in PGx-score research

While PGx-scores hold promise for predicting treatment outcomes, they currently account for only a small proportion of the variance in treatment outcomes. This systematic review highlights the lack of well-defined phenotypes and small samples sizes that limit our ability to adequately quantify the genetic complexity associated with medication response. In this context, the following future directions may improve the predictive capacity of PGx-score and move us closer to their clinical utilization in psychiatry.

Biologically informed PGx-scores: Previous PGx-score studies have been developed based on conventional polygenic modelling approaches, where the effect of genetic variants across the entire genome are aggregated, without taking into account the biological significance of these variants on the phenotype of interest. A biology-informed polygenic score (B-PGS) model was introduced very recently as a novel approach to improve both the predictive capability and biological meaning of polygenic scores, while also reducing sequencing costs. For example, in a study to predict psychosis, a pathway specific PGS that was restricted to genomic locations within “nervous system development” and “regulation of neuron differentiation”, explained a variance of 6.9% in the risk of psychosis, outperforming the conventional PGS where genome-wide SCZ variants accounted for only 3.7%. B-PGS potentially increase the polygenic signal to noise ratio by excluding variants with little association to pharmacogenomic outcomes and also enhance the clinical interpretability of polygenic models by focussing on specific molecular pathways. There is emerging evidence elsewhere in medicine that B-PGS may be useful for identification of new drug targets, for instance in inflammatory bowel disease.

Multi-trait PGx-score: By leveraging the genetic correlation between multiple phenotypes, the multi-trait PGS approach aggregates genetic information across traits with the aim to
improve the prediction power of PGx-scores \(^{101-103}\). For example, in patients with BD, the polygenic scores of SCZ or MDD explained 0.80% \(^{47}\) and 0.91% \(^{59}\) of the variance in lithium response, respectively. Interestingly, combining the polygenic scores of SCZ and MDD, resulted in a better model, with an explained variance of 1.85% in lithium treatment response \(^{41}\), indicating that multi-trait PGS outperform single trait PGS.

**Combining multimodal data and machine learning optimization:** Researchers have begun to combine PGS with other data modalities, for example clinical and imaging data to improve model accuracy \(^{40, 104}\). Machine learning methods are progressively being adopted for the analysis of multimodal or complex data comprising PGx-scores, socio-demographic, behavioural and clinical data \(^{105-106}\). This approach, exemplified in a few studies included in our review \(^{34, 40}\), holds promising results for clinical translation. Nevertheless, replication of these complex studies is lacking and interpretation of machine learning algorithms could be difficult for clinicians, potentially limiting their acceptance \(^{107, 108}\). To overcome this barrier, it is important for data scientists and clinicians to collaborate at an early stage of model development to ensure that these models are not only clinically useful but also calibrated and valid for local conditions and easily understandable for end users \(^{109-111}\).

**Validation of polygenic models:** Given the complexity of pharmacogenetic models, current sampling issues and the associated risks of false discovery and poor generalizability across different populations, external replication and validation of these models is critical for future implementation \(^{25, 112-114}\). Only 26.4% of studies included in this systematic review employed external validation \(^{31, 33, 34, 36, 37, 43, 45, 46, 55, 56, 59, 62, 64, 115}\).

**Multi-ancestry PGx-score:** Nearly 90% of samples in the target and discovery cohorts of studies included in our systematic review were European descent. Genetic variations and their effect on treatment outcomes can vary significantly among different populations. Given the
complex pattern of linkage disequilibrium (short genetic regions) and the significant difference in the frequency of genetic variants between populations, PGx-score constructed from one ancestral cohort may have lower prediction in another cohort. For instance in cardiovascular medicine, a Brazilian specific warfarin PGx-score used in a warfarin dosing algorithm was more accurate in Brazil, than the one developed in European population. Conversely, polygenic models that incorporate information from ancestrally diverse populations, improve prediction performance particularly in underrepresented non-European populations. Diverse sampling is required to develop and validate more generalizable and transferable PGx-scores across diverse populations. These limitations hamper the translation of research findings into clinical practice and raise health disparity concerns. Thus, improving diversity in pharmacogenomic research is essential steps in creating polygenic models with broader application.

**Clinical implications of PGx-score:** While it is clear that further development is required to improve accuracy of PGx-score and alone they have low clinical utility, findings are advancing our knowledge of pharmacogenomics toward better personalisation of treatment. For instance, the genetic loading for SCZ demonstrates some capability to stratify individuals based on lithium treatment response in BD and clozapine dosage in individuals with TRS. Drawing parallels from other disciplines, such as cardiovascular medicine, PGS for coronary artery disease been used to reclassify patients from intermediate into high-risk categories translating into stronger statin use recommendations. Similarly, genome-wide PGS in cardiovascular research have identified individuals with a four-fold increased risk, prompting recommendations for aggressive cholesterol-lowering therapy. Such evidence indicates that the polygenic scores have the potential to stratify patients, predict treatment outcomes and informed therapeutics decision making based on the genetic variation of population variation among different ancestral populations (Figure 4).
**Figure 4:** The potential use of PGx-score in precision psychiatry

*Abbreviations: DNA = Deoxyribonucleic acid*

**Conclusions**

In summary, this systematic review highlights that larger and more diverse target sample sizes, focussed on well-defined and standardised pharmacogenomic outcomes, with robust replication are required to optimise the development of PGx-scores. Currently the variance explained by these models is too small for effective clinical translation. However, new techniques, such as B-PGS and the use of multivariate modelling combining multiple traits PGS with clinical data look promising to the increase accuracy. Large scale consortia focused on pharmacogenomics are required to improve sample size and diversity.

**Funding**

AT Amare is currently supported by the National Health and Medical Research Council (NHMRC) Emerging Leadership (EL1) Investigator Grant (APP2008000). NT Sharew is a recipient of the University of Adelaide Research Scholarship.
References


4. WHO. Nearly one billion people have a mental disorders2022.


40. Cearns M, Amare AT, Schubert KO, Thalamuthu A, Frank J, Streit F *et al.* Using polygenic scores and clinical data for bipolar disorder patient stratification and lithium


46. Li J, Yoshikawa A, Brennan MD, Ramsey TL, Meltzer HY. Genetic predictors of antipsychotic response to lurasidone identified in a genome wide association study and by schizophrenia risk genes. Schizophr Res 2018; 192: 194-204.


48. Yoshida K, Marshe VS, Elsheikh SSM, Maciukiewicz M, Tiwari AK, Brandl EJ et al. Polygenic risk scores analyses of psychiatric and metabolic traits with antipsychotic-
induced weight gain in schizophrenia: an exploratory study. *Pharmacogenomics J* 2023.


