

Current level of evidence for improvement of antidepressant efficacy and tolerability by pharmacogenomic-guided treatment: A Systematic review and meta-analysis of randomized controlled clinical trials

Filip Milosavljević^a, Prof. Espen Molden^{b,c}, Prof. Magnus Ingelman-Sundberg^d,
Assoc. Prof. Marin M Jukić^{*,a,d}

^a Department of Physiology, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia

^b Center for Psychopharmacology, Diakonhjemmet Hospital, Oslo, Norway

^c Section for Pharmacology and Pharmaceutical Biosciences, Department of Pharmacy, University of Oslo, Oslo, Norway.

^d Pharmacogenetics Section, Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

ARTICLE INFO

Keywords:

Pharmacogenomics
Antidepressant
Major depressive disorder
CYP450
Psychiatry
Meta-analysis

ABSTRACT

The aim of the study was to assess the clinical utility of currently available pharmacogenomic (PGx) tools compared with treatment as usual (TAU), using a meta-analysis of dichotomous and continuous antidepressant efficacy and tolerability data from previously published clinical trials. MEDLINE, clinicaltrial.gov, EU Clinical Trials Register, WHO ICTRP and CENTRAL were systematically searched; of the 962 results originally reviewed, 15 trials were included. Antidepressant efficacy was quantified by relative and absolute changes in symptom severity after eight weeks of treatment and by response and remission rates, while tolerability was estimated by the rate of study discontinuation for any reason. In the PGx-guided patients, symptom severity reduced by an average of 31.0% after eight weeks of treatment, compared to an average reduction of 26.8% in the TAU group. Accordingly, PGx-guided patients experienced a greater reduction in symptom severity of 3.4% (95%CI: 1.6–5.3%), which corresponded to a reduction in the Hamilton Depression score of 0.75 (0.30–1.21), a 37% (15–63%) higher remission rate, and an 18% (5–33%) higher response rate compared with TAU patients, while no difference was observed in discontinuation rate between groups. Notably, the majority of associations lost statistical significance when restricting the dataset to low risk of bias studies, while certain funnel plots suggested a potential publication bias favoring the reporting of statistically significant results. In summary, PGx tools marginally enhance antidepressant efficacy, but not antidepressant tolerability; thus, additional research and advancement of PGx tools are needed to improve integration of PGx in clinical pharmacotherapy of depression.

1. Introduction

Pharmacogenomics (PGx) can be used to personalize and subsequently improve pharmacological treatment of major depressive disorder (MDD) by predicting and mitigating potentially harmful drug-gene interactions (Jukić et al., 2022). Research into numerous genetic biomarkers associated with antidepressant metabolism, efficacy and tolerability, including allelic variants of *CYP2C19*, *CYP2D6*, *SLC6A4*, *HTR2A*, *HLA-B* etc. (Jukić et al., 2022) has facilitated the development of several PGx-based clinical decision support tools aimed at improving antidepressant treatment outcomes (Bousman and Hopwood, 2016). Over the past decade, several randomized controlled clinical trials have been

conducted to investigate whether the use of the PGx tools developed to date can improve the efficacy and tolerability of antidepressant treatment compared with the treatment-as-usual (TAU) (Minelli et al., 2022, Zeier et al., 2018). However, the results of the individual trials were insufficient to draw a firm conclusion on this question (Bousman and Hopwood 2016; Minelli et al., 2022, Zeier et al., 2018, Zubenko et al., 2018, Goldberg et al., 2019a).

Recently, four meta-analyses (Brown et al., 2022, Arnone et al., 2023, Bunka et al., 2023, Wang et al., 2023) pooled the data from the available trials to obtain more meaningful results. All these meta-analyses detected a robust increase in remission and response rate as a result of PGx-guided treatment (Brown et al., 2022, Arnone et al.,

* Corresponding Author: Assoc. Prof. Marin M Jukić, Department of Physiology and Pharmacology, Karolinska Institute Solnavägen 9, B5 SE-171 65 Solna, Sweden.
E-mail address: marin.jukic@ki.se (Assoc.Prof.M.M. Jukić).

<https://doi.org/10.1016/j.euroneuro.2024.01.005>

Received 27 October 2023; Received in revised form 11 January 2024; Accepted 15 January 2024

Available online 9 February 2024

0924-977X/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2023, Bunka et al., 2023, Wang et al., 2023) and advocate clinical utility of currently available PGx tools for personalized antidepressant treatment of MDD. However, these meta-analyses (Brown et al., 2022, Arnone et al., 2023, Bunka et al., 2023, Wang et al., 2023) have mainly focused on dichotomous data such as response, remission and discontinuation rates, while the analysis of continuous outcomes was mainly absent. This is of major relevance because the percentage change from baseline in symptom severity at eight weeks has been reported as the primary endpoint in almost all available PGx trials and because it frequently occurred that certain efficacy outcomes within the same trial demonstrated the superiority of PGx-guided treatment over TAU, while others do not (Greden et al., 2019, Han et al., 2018, Perlis et al., 2020, Shan et al., 2019, Tiwari et al., 2022). Therefore, analyzing only a subset of outcomes does not provide the complete overview regarding clinical impact of PGx-guided treatment and continuous outcomes should also be investigated. Next, most of the available PGx trials presented outcome data for different time points after treatment initiation (Brown et al., 2022, Arnone et al., 2023, Bunka et al., 2023) and did not focus on a defined time point for all trials, which is of importance to standardize the conditions and minimize outcome heterogeneity. Finally, previous reports showed several methodological limitations: 1) Omission of several trials leading to the narrowing the scope of their meta-analyses, 2) the inclusion of the original trial data in the case where correction to the said data was published; 3) interchangeable use of Per protocol (PP) and Intention-to-treat (ITT) data; and 4) Not exploring demographics and methodological heterogeneity in sensitivity analysis and only using “leave-one-out” method (Table 1). In conclusion, the listed shortcomings argue for a need for a more comprehensive meta-analysis on this topic.

The aim of this systematic review and meta-analysis was to build on recent work and comprehensively describe all available data, both dichotomous and continuous, from the representative PGx trials. This analysis aims to clarify whether antidepressant therapy for patients with MDD, guided by the currently available PGx tools, leads to improvements in efficacy and tolerability compared to treatment as usual, and whether these effects are statistically and clinically significant.

2. Materials and methods

This report was written in accordance with the PRISMA 2020 guidelines (Page et al., 2021). The protocol for this meta-analysis series was not pre-registered because the results of the literature search were originally taken from another research project (Jukić et al., 2022) and the authors detected no additional benefit in registering the protocol after the original literature search that was published.

2.1. Literature search and study selection

Several digital databases were searched including: MEDLINE, clinicaltrials.gov, EU Clinical Trials Register, WHO ICTRP and CENTRAL. Published and unpublished reports of clinical trials conducted between January 1990 and 19th September 2022 (most recent re-run 24th October 2023) were screened for the inclusion. The search terms used were: "(personaliz* OR pharmacogen*) AND depress* AND randomized" (MEDLINE, EU Clinical Trials Register and WHO ICTRP), "Pharmacogenomics Depression Randomized Controlled Trial" (CENTRAL) and "Randomized Depression Pharmacogenomics" (clinicaltrials.gov).

A clinical trial was included if: (i) Participants were randomized between the TAU and PGx-guided groups; (ii) participants were treated with antidepressants for MDD; and (iii) MDD symptom severity was assessed using a validated clinical assessment scale; specifically, the Hamilton Depression Rating Scale (HAM-D), the Montgomery-Asberg Depression Rating Scale (MADRS), the Patient Health Questionnaire 9 (PHQ-9), the Quick Inventory for Depressive Symptomatology (QIDS) or the Children's Depression Rating Scale (CDRS). Trials with serious methodological flaws were excluded post-hoc. No restrictions were made on manuscript language, age of participants, type of

antidepressant prescribed or type of PGx tool used. Searching, screening, data extraction, and transformation were performed by FM. The final selection of trials was decided by consensus between FM and MMJ. Risk of bias was assessed using the ROB2 tool (Sterne et al., 2019) across five domains.

2.2. Data extraction

All outcome data measured eight weeks after randomization were extracted; if no measurements were taken exactly after eight weeks, the measurement closest to the eight weeks within the interval of 6-12 weeks was selected. Extracted outcome data for the TAU and PGx arms included: (i) Symptom severity at baseline and at eight weeks; (ii) Nominal and relative change in symptom severity at eight weeks; (iii) Response rate at eight weeks, where response was defined as a reduction in symptom severity greater than 50%; (iv) Remission rate at eight weeks, where remission was defined as a symptom severity level HAM-D ≤ 7 , CGI-I ≤ 2 , PHQ ≤ 5 , QIDS-SR16 ≤ 6 ; CDRS-R ≤ 28 or MADRS ≤ 10 ; (v) Rate of discontinuation for any reason at eight weeks; and (vi) Frequency, intensity, burden of adverse events (FIBSER) frequency score at baseline and eight weeks. For continuous outcomes, the means and standard deviations were extracted for both arms, while for dichotomous outcomes, the total number of affected and unaffected participants per arm was extracted.

When available, both per-protocol (PP) and intention-to-treat (ITT) data were extracted. ITT data were prioritized and data from PP were only included if ITT data were not available. As most of the included trials had published both ITT and PP data for dichotomous endpoints (response and remission rates), the differences between these approaches were explored in the sensitivity analyses and presented in the supplement. The exact definitions of the endpoints selected for each meta-analysis conducted can be found in the supplement (Table S3). The following data were extracted from each included trial for additional analyses: (i) patient demographics (male/female ratio, mean age, ethnic background, baseline symptom severity), (ii) patient inclusion/exclusion criteria, (iii) PGx tool characteristics, (iv) antidepressants prescribed, (v) funding sources, (vi) type of allocation blinding, (vii) choice of primary and secondary endpoint, (viii) protocol pre-registration, (ix) total duration of follow-up and (x) data required for risk of bias analysis, i.e. the data on blinding, deviation from the intended intervention, absence of data, outcome measurement and data reporting.

2.3. Missing data imputation procedures

If nominal change in symptom severity at eight weeks was not reported, but values for symptom severity at baseline and week eight were available, this outcome was estimated using a previously published method (Higgins et al., 2022). If the percentage change in symptom severity between baseline and week eight was not reported, but baseline symptom severity and nominal symptom severity change were reported or estimated for the same time points, the Taylor expansion method was used to estimate this outcome (Stuart and Ord, 1998). Standard errors and confidence intervals were converted to standard deviations using standard formulae (Higgins et al., 2022), while data shown graphically were extracted by extrapolating the data shown in the figure. Nominal changes in symptom severity scores at eight weeks, as measured by rating scales in the included trials, were converted to nominal changes in the HAM-D scale for the same time points, based on previously published correlation formulae (Sun et al., 2020, Jarrett et al., 2008). When remission and response rates were not reported by the authors, but continuous depression severity data were available, response and remission rates were estimated using the method described previously (Furukawa et al., 2005). The validity of the transformations was ascertained in previous papers (Weir et al., 2018; Kambach et al., 2020) and double-checked by using data from certain trials (Han et al., 2018, Oslin et al., 2022, Perez et al., 2017) where both input and output data for the

Table 1

Comparison of inclusion of outcome data and summary results between the current and previous meta-analyses. Differences in previous meta-analyses in the selection of input data compared to the current report are highlighted in red. Regarding the meta-analyses on remission and response rates, the current report provides the largest scope compared to the four previous meta-analyses, but still excludes study with critical methodological problems (Bradley et al., 2018). The meta-analysis was performed using the "Mantel-Haenszel" method and a random effects model.

RCT	Outcome	Current meta-analysis	Brown et al. 2022 ¹	Arnone et al. 2023	Bunka et al. 2023	Wang et al. 2023
Bradley et al., 2018	Response rate	Not analyzed	Not analyzed	Week 12, PP Cohort subset (unclear)	Week 12, PP, Cohort subset with HAM-D \geq 25	Week 8, PP, Cohort subset with HAM-D \geq 20
	Remission rate	Not analyzed	Week 12, PP, Cohort subset with HAMD \geq 25	Week 12, PP Cohort subsets with HAM-D \geq 25	Week 12, PP, Cohort subset with HAM-D \geq 25	Week 8, PP, Cohort subsets with HAM-D \geq 25
Greden et al., 2019	Response rate	Week 8, ITT	Unclear	Week 8, PP	Week 8, ITT	Week 8, ITT ²
	Remission rate					
Han et al., 2018	Response rate	Week 8, ITT Corrected version of the manuscript	Week 8, ITT Uncorrected version of the manuscript	Week 8, ITT Corrected version of the manuscript	Week 8, ITT Uncorrected version of the manuscript	Week 8, ITT Uncorrected version of the manuscript
	Remission rate					
McCarthy et al., 2021	Remission rate	Week 8, ITT	Week 8, PP	Week 8, PP	Not analyzed	Unclear
Oslin et al., 2022	Response rate	Week 8, ITT	Week 24 ITT	Week 24, PP	Not analyzed	Week 8, ITT
	Remission rate					
Papastergiou et al., 2021	Response rate	Week 12, PP	Not analyzed	Not analyzed	Not analyzed	Not analyzed
	Remission rate					
Perez et al., 2017	Response rate	Week 12, ITT, HAM-D defined	Not analyzed	Week 12, PP, PGI-I defined	Unclear, HAM-D defined	Not analyzed
	Remission rate	Week 12, ITT, HAM-D Defined	Week 12, PP HAM-D Defined	Not analyzed	Unclear	Not analyzed
Perlis et al., 2020	Response rate	Week 8, ITT	Week 8, PP	Week 8, PP	Week 8, PP	Week 8, PP
	Remission rate	Week 8, ITT	Week 8, PP	Week 8, PP	Week 8, PP	Not analyzed
Shan et al., 2019	Response rate	Week 8, ITT	Week 8, ITT	Week 8, ITT	Week 8, ITT	Week 8, ITT
	Remission rate					
Singh 2015	Remission rate	Week 12, PP	Week 12, PP	Week 12, PP	Week 12, PP	Not analyzed
Tiwari et al., 2022	Response rate	Week 8, ITT	Unclear	Week 8, PP	Not analyzed	Week 8, ITT
	Remission rate					
Vande Vroot et al., 2021	Response rate	Week 8, ITT	Not analyzed	Not analyzed	Not analyzed	Not analyzed
	Remission rate					
Vos et al., 2023	Remission rate	Week 7 ITT	Not analyzed	Not analyzed	Not analyzed	Not analyzed
Winner et al., 2013	Response rate	Week 10, ITT	Week 10, PP	Week 10, PP	Week 10, PP	Week 10, PP
	Remission rate					

Summary results across different meta-analysis (Grand mean / [95%CI] / Total number of patients / Significance level)

(continued on next page)

Table 1 (continued)

RCT	Current meta-analysis		Brown et al. 2022 ¹		Arnone et al. 2023		Bunka et al. 2023		Wang et al. 2023							
	Outcome	Risk ratio	N	p	N/A	N/A	N/A	N/A	N/A	N/A						
Response rate	1.18	[1.05, 1.33]	N=4779	p=0.0044	N/A	N/A	N/A	1.32	[1.00, 1.73]	N=2188	p=0.047	1.21	[1.06, 1.40]	N=4328	p=0.0061	
Odds ratio	1.30	[1.08, 1.55]	N=4779	p=0.0045	N/A	N/A	1.46	[1.16, 1.85]	1.68	[1.11, 2.54]	N=2188	p=0.015	1.34	[1.07, 1.69]	N=4328	p=0.012
Remission rate	1.37	[1.15, 1.63]	N=5201	p=0.0004	1.41	[1.15, 1.74]	N=4767	p=0.0012	N/A	N/A	N=2341	p=0.043	1.46	[1.25, 1.71]	N=4071	p<0.0001
Odds ratio	1.56	[1.19, 2.04]	N=5201	p=0.0011	1.70	[1.24, 2.34]	N=4767	p=0.0010	1.85	[1.23, 2.76]	N=2341	p=0.014	1.58	[1.30, 1.91]	N=4071	p<0.0001

¹ The authors did not analyze remission rates and also included 2 non-randomized studies not included in other meta-analyses (not shown). [†]Data for remission rates are estimated, although original data are available. RCT: Randomized controlled trial; PP: Per protocol; ITT: Intention-to-treat; HAM-D: Hamilton depression rating scale; PGI-I: Patient's global scale of impression for treatment improvement; 95%CI: 95% Confidence interval; N/A not applicable.

respective outcome transformation process had already been provided in respective reports. If the published trial data were not sufficient for inclusion even after all listed procedures, the corresponding author was contacted to provide additional data; subsequently, a trial was excluded if the data were ultimately not sufficient for the analysis of a specific outcome, one month after the contact attempt.

2.4. Meta-analyses

RevMan V5.4.1 software (Cochrane Collaboration, 2020) was used for the meta-analyses. Dichotomous outcomes were analyzed using the risk ratio metric and the Mantel-Haenszel method, while the odds ratio metric was calculated only for the purpose of comparison with other meta-analyses. Comparison between groups for continuous outcomes was performed using the mean difference metric and the inverse variance method. Due to the inherent heterogeneity resulting from the different PGx tool and choice of antidepressants in the included trials, a random effects model was used for all meta-analyses. The percentage change from baseline in symptom severity at eight weeks was the primary meta-analysis for antidepressant efficacy. Three additional secondary meta-analyses of antidepressant efficacy were conducted: (i) meta-analysis of nominal change from baseline in HAM-D scores; (ii) meta-analysis of response rate; and (iii) meta-analysis of remission rate. Due to the limited availability of data on antidepressant adverse drug reactions, no reliable meta-analysis on antidepressant tolerability could be performed. However, we conducted a meta-analysis on the discontinuation rate for any reason, supplemented by the secondary meta-analysis of the change from baseline in FIBSER frequency scores. A sensitivity analysis was performed to examine the robustness of the effect to changes in the trial inclusion and exclusion criteria; inclusion was restricted to trials: (1) associated with low risk of bias, (2) with more than hundred patients, (3) dealing with adult patients (18-65 years old) only, (4) with outcome measured exactly eight weeks after treatment initiation, (5) with no missing data imputation needed. Also, exploratory analyses included the analysis where the trial with the most pronounced positive result was excluded; inclusion was limited to Caucasians or East-Asians alone, and inclusion was limited to usage of a distinct PGx tool. Small-study effect and potential publication bias were assessed using visual inspection of contoured funnel plots (Peters et al., 2008). Sterne et al. (2019) and Egger's test (Egger et al.,1997) and the methodology behind these tests is described in full detail in the supplement.

2.5. Interpretation of clinical meaningfulness

To gauge the clinical meaningfulness of presently available PGx tools based on continuous outcomes the effect was compared with a previously established criterion that the lowest observable change in symptom severity reduction is a three-point reduction in HAM-D score (Leucht et al., 2013). Evaluation of the clinical meaningfulness of PGx tools based on categorical outcomes involved the use of number needed to treat (NNT) and number needed to genotype (NNG) metrics for response and remission rates. In essence, NNT and NNG, in the context of this study, represents the number of individuals requiring treatment or genotyping to facilitate one additional event of remission or response. Calculated based on event risks among PGx-guided individuals and those treated-as-usual (Tonk et al., 2017), values lower than 10 are considered clinically meaningful, aligning with previously established criteria (Goldberg et al., 2019a).

3. Results

3.1. Summary of the trial and participant characteristics

Of the 962 records reviewed, 15 trials met the inclusion criteria (Bradley et al., 2015, Greden et al., 2019, Han et al., 2018, McCarthy et al., 2021, Oslin et al., 2022, Papastergiou et al., 2021, Pérez et al.,

2017, Perlis et al., 2020, Shan et al., 2019, Singh 2015, Tiwari et al., 2022, Vande Vroot et al., 2021, van der Schans et al., 2019, Vos et al., 2023, Winner et al., 2013) (Fig. 1). The cohort characteristics (Table S1) and study design (Table S2) of the included trials are presented in detail in the supplement. Overall, most trials included predominantly female and predominantly Caucasian patients with a mean age of around 45 years. Exceptions were: one trial with adolescents (Vande Vroot et al., 2021), one trial with older patients (van der Schans et al., 2019), and two trials with East Asian patients (Han et al., 2018, Shan et al., 2019). Six trials included patients who had a history of psychiatric drug failure, three included patients taking an antidepressant for the first time, three included both categories, and in three trials this was unclear. All but one study [28] did not restrict the choice of antidepressant, while most trials allowed the introduction of the second antidepressant. The most commonly used PGx tool was GeneSight® (Greden et al., 2019, Tiwari et al., 2022, Oslin et al., 2022, Vande Vroot et al., 2021, Winner et al., 2013), while other PGx tools used were: Neuropharmagen® (Han et al., 2018, Pérez et al., 2017), PillCheck® (Papastergiou et al., 2021), NeuroIDgenetix® (Bradley et al., 2018), Genecept (Perlis et al., 2020), CNSdose® (Singh 2015); Mental Health DNA Insight (McCarthy et al.

2021); and three trials used non-commercial PGx solutions and protocols (Shan et al., 2019, Vos et al., 2023, van der Schans et al., 2019). The duration of follow-up ranged from eight weeks to six months. Due to the nature of reporting PGx results in the trials, it was virtually impossible to blind treating physicians to patient allocation in an ethical way. Therefore, most of the included studies used blinded clinical raters who independently scored the clinical scales and all trials except Oslin et al. 2022 also blinded the patients.

Two of the included trials (Bradley et al., 2018, Singh, 2015) had critical methodological problems and were therefore excluded from certain analyses. In particular, the study by Bradley et al., 2018 reported discontinuation rates for the entire cohort, while data on changes in HAM-D scores were only reported for those patients with baseline HAM-D scores of more than 20 points, but not for the other patients. Because the authors selectively reported efficacy data only for the subgroup of the randomized cohort while omitting the other patients, and because the corresponding author never responded to multiple email inquiries, this trial was removed from all efficacy analyses because of the high risk of reporting bias. Singh 2015 trial reported highly imbalanced baseline depression severity scores between PGx and TAU patients. As

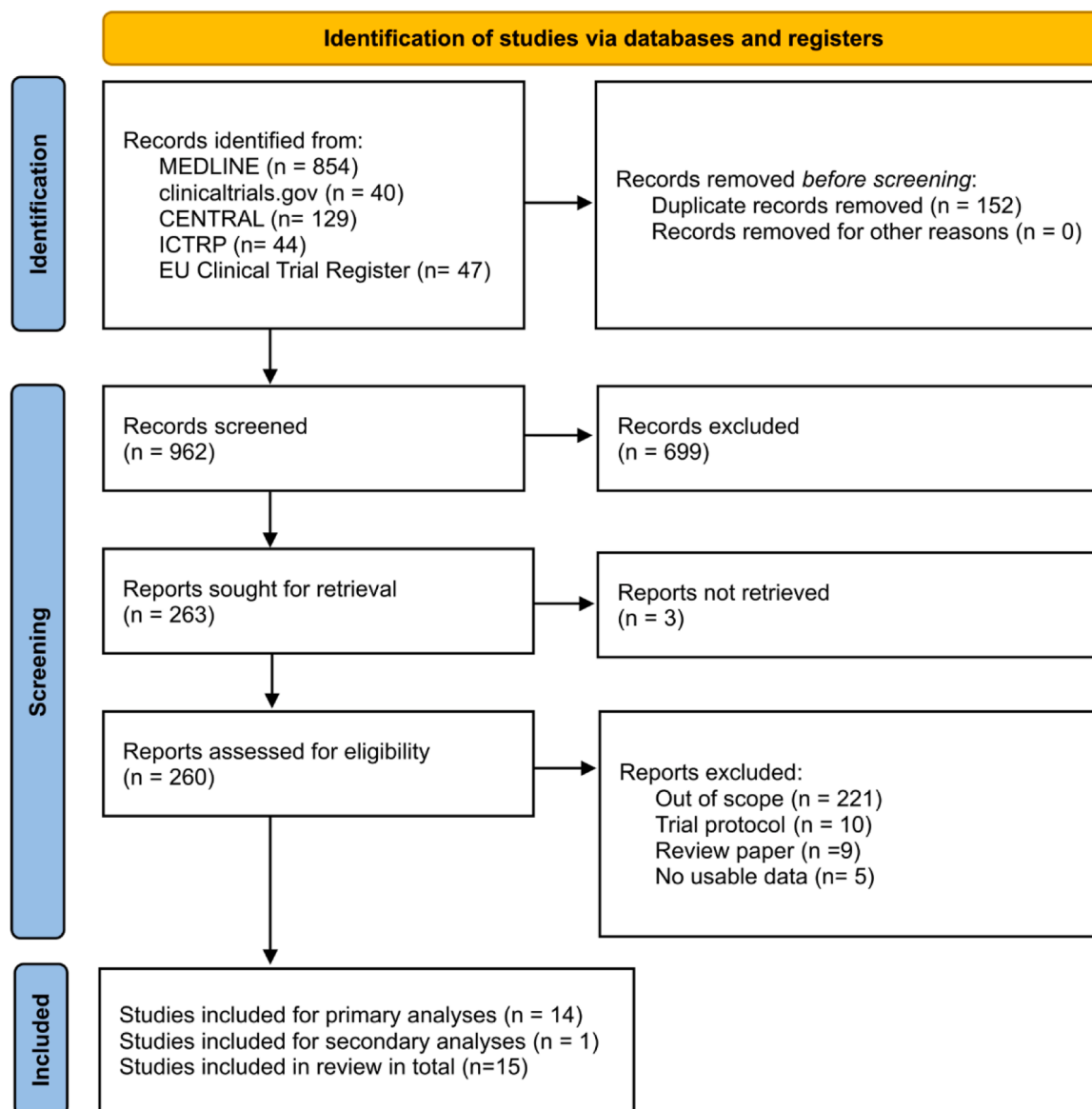


Fig. 1. PRISMA flowchart for literature search: Out of 962 screened records from several digital literature databases, 15 RCTs were included across different meta-analyses. The flowchart is representative of the most recent literature search re-run performed on October 24th 2023.

higher baseline severity leads to higher nominal and relative reductions in HAM-D score during follow-up (Hieronymus et al., 2019), this study was only included in the analysis of remission and discontinuation rates. Additionally, due to the age being an important factor in antidepressant treatment, trials on adolescents (Vande Vroot et al., 2021) and elderly cohort (van der Schans et al., 2019) were excluded from the calculations as a part of sensitivity analysis. Using the RoB2 analysis of risk of bias (Table S9), five trials were rated as having a low risk of bias, seven trials were rated as having a serious risk of bias, and three trials were rated as having a high risk of bias. The most common causes of risk of bias were suboptimal handling of missing information, lack of a preregistered study protocol and suboptimal concealment of randomization.

3.2. The effect of pharmacogenomic guidance on antidepressant efficacy and tolerability

The primary endpoint of antidepressant efficacy, the percentage change in MDD symptom severity after eight weeks of antidepressant treatment (Fig. 2A), showed a mean reduction in symptom severity of 31.0% (95%CI: 29.7, 32.3%) in the PGx-guided group and a mean reduction of 26.8% (95%CI: 25.6–28.1%) in the TAU group; comparison of these values resulted in a statistically significant 3.4% (95%CI: 1.6%–5.2%) better improvement in PGx-guided compared to TAU group. When comparing the nominal reduction in HAM-D score after eight weeks from baseline (Fig. 2B), symptom improvement was 0.75 HAM-D points higher in the PGx-guided than in the TAU group (95%CI: 0.30–1.21). Meta-analyses of dichotomous outcomes related to antidepressant efficacy showed a mean response rate of 29.4% (95%CI: 27.6%–31.2%) in the PGx group and 24.8% (95%CI: 23.0%–26.6%) in the TAU group, which translated into 1.18 times (95%: 1.05–1.33) higher response rate in the PGx-guided compared to TAU group (Fig. 2C). In addition, the mean remission rate was 19.5% (95%CI: 18.1%–21.0%) in the PGx-guided group and 14.7% (95%CI: 13.3%–16.2%) in the TAU group, which translated into 1.37-fold (95%CI: 1.15–1.63) higher probability of remission in the PGx-guided compared to TAU group (Fig. 2D). The meta-analysis of remission rate showed moderate heterogeneity ($I^2=50\%$), which can be attributed to Perlis et al. 2020 and Singh 2015, as the exclusion of either of these two trials abolishes the heterogeneity within this meta-analysis.

There were no significant differences between the PGx and TAU groups in the rate of treatment discontinuation for any reason after eight weeks (Fig. 3A). The meta-analysis on FIBSER frequency score showed no differences between the PGx and TAU arms in the change from baseline (Fig. 3B). Meta-analysis of change from baseline in FIBSER frequency score showed significant heterogeneity ($I^2=57\%$), likely due to the small number of included trials. In summary, the use of currently available pharmacogenomic tools statistically significantly improves efficacy with no impact on discontinuation rate.

3.3. Effect Robustness and Consideration of Publication Bias

The impact on the relative change in symptom severity from baseline lost significance in the sensitivity analysis when the inclusion was limited exclusively to trials with low risk-of-bias (Table S4), while the effect on the nominal reduction in HAM-D scores maintained significance across all sensitivity analyses (Table S5). Notably, a substantial alteration in effect magnitude emerged in the sensitivity analysis involving only East Asian patients (Shan et al., 2019; Han et al., 2018). In this case, the improvement in symptom severity reduction within the PGx-guided group compared to the treatment-as-usual (TAU) group escalated significantly to 10.6% relative change or a nominal 2.8 HAM-D decrease. The sensitivity analyses focused on response and remission rates revealed a loss of statistical significance for both effects when inclusion was restricted exclusively to trials with low risk-of-bias (Table S6 and S7). Additionally, the effect on response rate was no longer

significant when inclusion was limited exclusively to large trials (refer to Table S6). Individual trials exhibited notable influence on the meta-analysis of remission rates; specifically, excluding the Singh 2015 study decreased the risk ratio from 1.37 to 1.27, while exclusion of the Perlis et al. 2020 study increased the risk ratio from 1.37 to 1.45 (refer to Table S7). As it was the case for the continuous endpoints, the risk ratio of response rates favouring the PGx group increased from 1.18 to 1.42 when the inclusion was limited to East Asians exclusively, while the difference in remission rates remained consistent. In an exploratory sensitivity analysis focusing on dropout rates, certain changes to inclusion criteria resulted in a significant advantage favouring better tolerance in the TAU group compared to the PGx arm. Concerning the use of individual tools, restricting the analysis to studies using Genesight® PGx tools yielded results consistent with the principal meta-analyses, while sub-analyses of other PGx tools remained inconclusive due to limited data availability.

Visual assessment of the contoured funnel suggests asymmetry due to the possibility of missing studies with low precision and statistically non-significant results indicates asymmetry for the relative change in symptom severity (Fig. 4A), nominal change in symptom severity (Fig. 4B), and response rate (Fig. 4D). The statistically significant in Egger's test for the change in HAM-D score from baseline (Fig. 4B) indicated that publication bias cannot be ruled out for this effect. While the significant asymmetry for the discontinuation rate that was also observed in Egger test (Fig. 4C), this asymmetry was likely caused by methodological heterogeneity between studies rather than by publication bias because the funnel plot clearly indicates that the asymmetry is not related to the significance level of the outcome (Sterne et al., 2019). In the meta-analyses of remission rate (Fig. 4E), no significant asymmetry was found either visually or statistically. In summary, the robustness of the observed effect is not unequivocal, as the analysis of studies associated with low risk of bias does not indicate superiority PGx-guided treatment compared to TAU, while publication bias cannot be equivocally ruled out for most of the analyzed outcomes.

3.4. Clinical meaningfulness of the effect

Related to continuous outcomes, the difference in HAMD score reduction between PGx and TAU arms after eight weeks was 0.75 (95% CI: 0.30–1.21), which is below the level regarded as clinically meaningful according to previously established criteria (Leucht et al. 2013). Related to categorical outcomes, the number of patients needed to treat (NTT) to achieve one additional response was 21.1, while the NNT to achieve one additional remission was 21.3, meaning that both were substantially higher than 10, which is a threshold for clinically meaningful effect according to the previously established criterion (Goldberg et al., 2019a). Insufficient data was presented for number needed to genotype metric calculations, as the response and remission rates for participants with variant and reference alleles within PGx-guided groups was seldom reported separately. In summary, even with concerns related to the effect robustness and publication bias aside, the effect size based on the outcomes indicating the improvement of treatment is below the threshold to be considered clinically meaningful.

4. Discussion

To our knowledge, this manuscript provides the most comprehensive and balanced meta-analysis studying the impact of antidepressants on both continuous and dichotomous efficacy parameters in clinical trials comparing standard treatment and treatment guided by currently available pharmacogenomic clinical decision support tools in patients with major depression. Although the results suggest a statistically significant increase in antidepressant efficacy, a possible bias towards positive results, the small size of the effect, and the insufficient robustness of the observed increase, together with the lack of impact on antidepressant tolerability, suggest that the clinical utility of currently

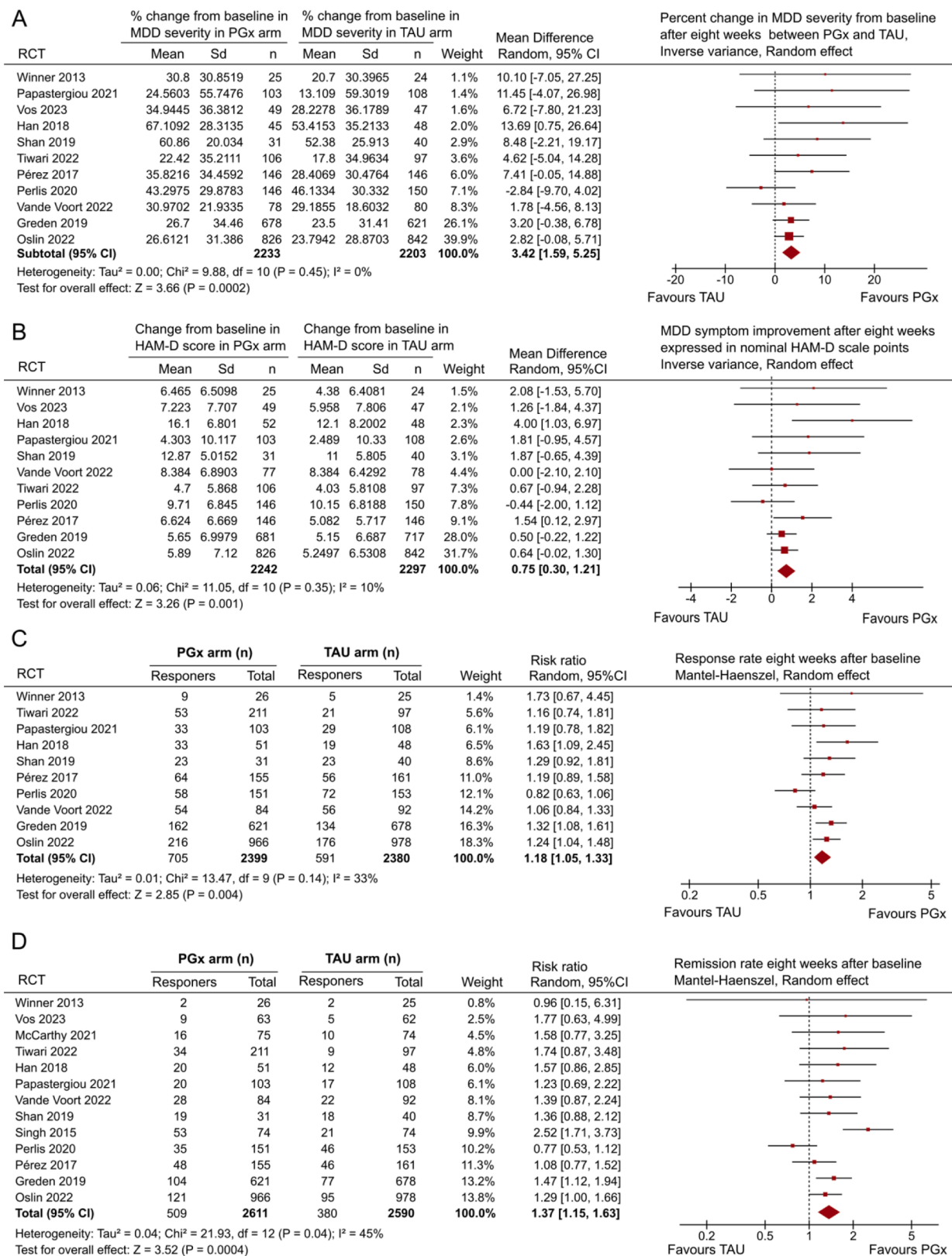


Fig. 2. Forest plots of antidepressant efficacy comparison between PGx and TAU arms. (A) The PGx-guided arm had an average reduction in symptom severity of 31.0% (95%CI: 29.7, 32.3%), while the TAU group had a mean reduction in symptom severity of 26.8% (95%CI: 25.6, 28.1%); thus, the mean difference was 3.42% (95%CI: 1.59, 5.25%; df=10, p=0.0002, n=11, N=4 436). (B) After eight weeks of MDD treatment, PGx guided arm had on average 0.75 HAM-D points (95%CI: 0.30, 1.21; df=10, p=0.0011, n=11, N=4 539) better treatment improvement as compared to TAU arm. After eight weeks of MDD treatment, (C) there was 18% (95%CI: 5%, 33%; df=9, p=0.0044, n=10, N=4 779) higher probability of achieving treatment response and (D) 37% (95%CI: 15%, 63%; df=12, p=0.0004; n=13, N=5 201) higher probability of achieving remission in PGx guided arm compared to TAU arm. PGx guided: Pharmacogenomics guided therapy of depression; TAU: Treatment-as-usual; MDD: Major depressive disorder; RCT: Randomized controlled trial; CI: Confidence interval; SD: Standard deviation.

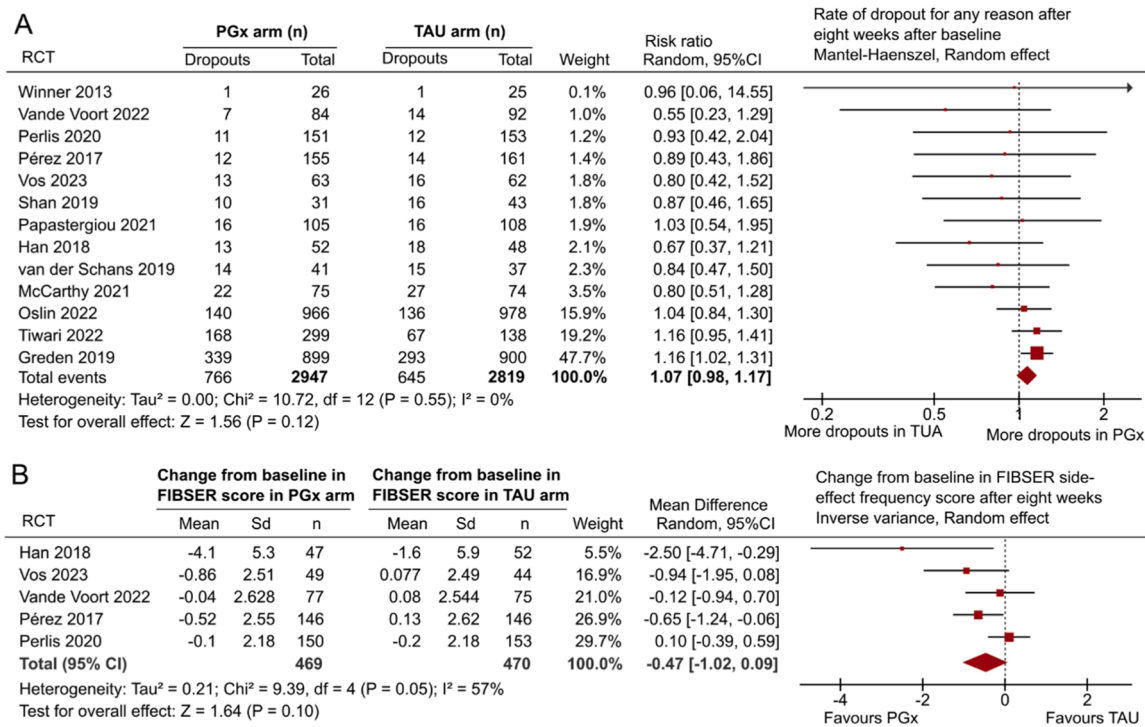


Fig. 3. Forest plots of antidepressant tolerability comparison between PGx and TAU arms. (A) There were no significant differences in discontinuation rates between the PGx-guided and TAU arm (Risk ratio 95%CI: 0.98, 1.17; df=12, p=0.12, n=13, N=5 766). (B) There was no significant difference in FIBSER frequency score (item 1) change from baseline until week eight (Mean difference: -0.47; 95%CI: -1.02, 0.09; df=4, p=0.10, n=5, N=939) of treatment between PGx guided and TAU arms. PGx-guided: Pharmacogenomics guided therapy of depression; TAU: Treatment-as-usual; RCT: Randomized controlled trial; CI: Confidence interval; SD: Standard deviation; HAM-D: Hamilton depression rating scale; FIBSER: Frequency, intensity, burden of side-effects rating scale.

available pharmacogenomic tools in antidepressant treatment is still questionable.

Four previously published meta-analyses (Brown et al., 2022, Arnone et al., 2023, Bunka et al., 2023, Wang et al., 2023) had a similar aim like this analysis, but then focused only on dichotomous outcome measures, which were defined as secondary endpoints in the included trials. A direct comparison between these analyses showed many similarities, such as the fold-change in remission rate between groups, suggesting that one additional remission event per twenty patients guided with PGx is expected. However, there are also many inconsistencies due to differences in methodology and interpretation (Table 1). Here, these inconsistencies through an umbrella review of previously published reports and re-analysis of the data; specifically, (i) the trials with serious methodological problems were excluded, (ii) the time at which the results were analyzed was standardized, and (iii) the results were analyzed primarily using the intention-to-treat approach rather than the per protocol approach, in line with the well-established practice (Samara et al., 2016). The present analysis shows a less pronounced increase in response and remission rates in the PGx compared to TAU group and suggests a lack of robustness (Table 1). Furthermore, while response and remission rates are highly relevant, the threshold-based binary categorization results in a loss of differences in response intensity between patients and limits the quantitative assessment of the magnitude of the effect of the PGx tool usage on antidepressant efficacy. Here, we focused on both dichotomous and continuous outcomes and were able to quantify a nominal 3.5% greater reduction in symptom severity from baseline eight weeks after treatment initiation, corresponding to an average improvement of 0.77 HAM-D points in the PGx group compared to TAU. Clinical significance of these results can be assessed in two ways: 1) Changes in HAM-D score were lower than 3 points are considered clinically meaningful (Leucht et al., 2013); and 2) Number needed to treat (NTT) metric for both remission and response rates for our meta-analysis were 21 while, while NTT less than 10 is considered clinically meaningful (Goldberg et al., 2019a). In other words, 21

patients need to be treated with either currently available PGx tool in order to produce one additional MMD remission or response. Thus, it is difficult to conclude unequivocally that the observed effects on currently available PGx tools are clinically relevant given the small effect size and the possibility of the publication bias.

Importantly, all included trials were either inconclusive or negative stand-alone; specifically, they either did not have a preregistered primary endpoint or the preregistered primary endpoint did not show superiority of the PGx tool compared to TAU. Pre-specification of the primary endpoint and adherence to this endpoint later in the analysis and reporting is a very important practice that reduces unconscious and conscious bias, while without this, it is difficult to assert causality between independent and dependent variable (Goldberg 2019b). Examples of this are the trials published by Oslin et al., 2022 and Papastergiou et al. 2021, which reported positive results, but included significant discrepancies between the predefined protocol and the published results (Table S2).

Finally, the true potential of PGx to improve antidepressant treatment outcomes has likely been diluted by a substantial proportion of placebo responders and absolute non-responders (Preskorn, 2014) and attenuated by inherent shortcomings of currently available PGx tools. In particular, currently available PGx tools sometimes include variants whose relevance to antidepressant treatment are not yet clearly established and fully understood (Border et al., 2019) while omitting clinically relevant genetic variants and diplotypes, such as CYP2C19 and CYP2D6 loss-of-function alleles (Jukić et al., 2022). Moreover, these PGx tools usually translate genetic biomarkers into clinical recommendations in a proprietary and non-transparent manner. Further, rather than giving the precise treatment guidelines, most available PGx tools serve to alert clinicians to avoid potentially relevant gene-drug interaction, which may lead to the unneeded discontinuation of potentially effective drugs (Goldberg et al., 2019a). In addition, the included trials were heterogeneous in terms of choice of PGx tool, patient cohort and type of antidepressants used (Table S2). Although the results of sensitivity analysis presented here may provide additional context (Tables S4–S8),

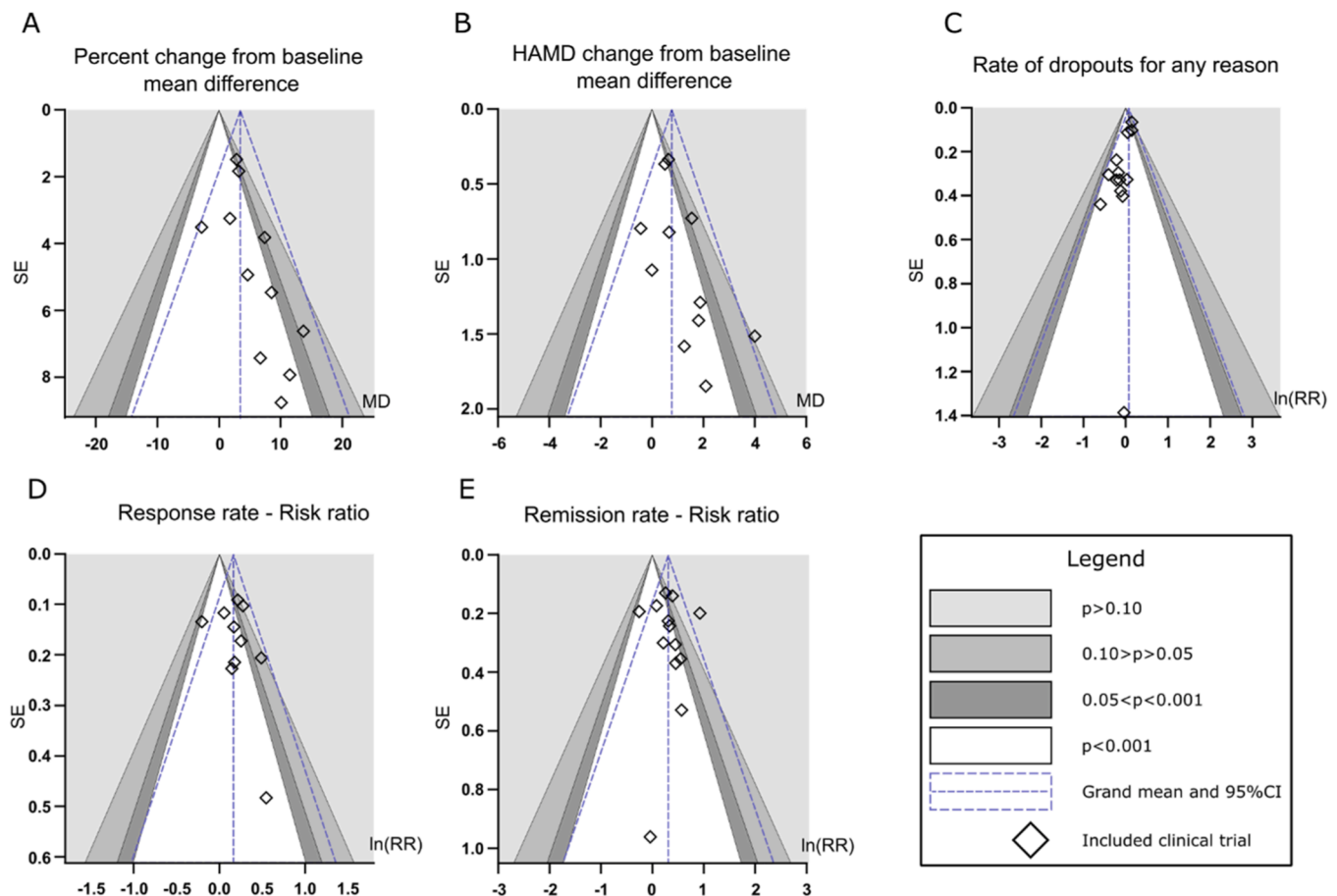


Fig. 4. Contour-enhanced funnel plots: Visual assessment of contour-enhanced funnel plots revealed possible asymmetry in (D) response rate ($p=0.67$) while both visual and statistically significant asymmetries were observed for (A) percentage change from baseline ($p=0.061$), (B) nominal change from baseline in HAM-D score ($p=0.038$) and (C) discontinuation rate for any reason ($p=0.0002$) meta-analyses. No asymmetry was observed for (E) remission rate ($p=0.11$) meta-analyses. Significance of the funnel plot asymmetry was tested using Egger’s test, where $p < 0.10$ is considered statistically significant [16]. SE: Standard error; MD: Mean difference; HAM-D: Hamilton depression rating scale; RR: Risk ratio; CI: Confidence interval

our summary results may have limited generalizability because the method we used is inherently unable to account for existing heterogeneities. Finally, although data imputation methods show more correct results compared to the omission of the incomplete datasets (Weir et al., 2018, Kambach et al., 2020), it is impossible to completely eliminate the risk of inaccurate data estimation in our meta-analyses.

In summary, adequately exploiting the true potential of pharmacogenomic biomarkers requires further genetic research focusing on individual gene-drug interactions, better defined interventions based on individual drug-biomarker interactions, and sufficiently powered and designed clinical trials focusing on such improved and standardized pharmacogenomic clinical decision support tools.

Declaration of competing interest

The authors declare no conflict of interest.

Role of funding source

This research was funded by the Swedish Research Council (grant 2021-02732), the Science Fund of the Republic of Serbia (grant 6066800), and the Swedish Brain Foundation (grants FO2021-0314 and FP2023-0189). Funding source had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Contributors

All authors significantly participated in the conceptualization and study design; FM conducted the data collection; FM and MMJ conducted the data analysis; MMJ, MIS and EM interpreted the analyzed data. FM and MMJ drafted the manuscript, MIS and EM critically revised it for important intellectual content. All authors approved the final version of the manuscript for the submission. All authors agreed to be accountable for all aspects of the work and to ensure that issues related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgement

This research was funded by the Swedish Research Council (grant 2021-02732), the Science Fund of the Republic of Serbia (grant 6066800), the Swedish Brain Foundation (grant FO2021-0314).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.euroneuro.2024.01.005](https://doi.org/10.1016/j.euroneuro.2024.01.005).

References

Arnone, D., Omar, O., Arora, T., Ostlundh, L., Ramaraj, R., Javaid, S., Govender, R.D., Ali, B.R., Patrinos, G.P., Young, A.H., Stip, E., 2023. Effectiveness of pharmacogenomic tests including CYP2D6 and CYP2C19 genomic variants for

- guiding the treatment of depressive disorders: Systematic review and meta-analysis of randomised controlled trials. *Neurosci Biobehav Rev* 144, 104965.
- Border, R., Johnson, E.C., Evans, L.M., Smolen, A., Berley, N., Sullivan, P.F., Keller, M.C., 2019. No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples. *Am J Psychiatry* 176 (5), 376–387.
- Bousman, C.A., Hopwood, M., 2016. Commercial pharmacogenetic-based decision-support tools in psychiatry. *Lancet Psychiatry* 3 (6), 585–590.
- Bradley, P., Shiekh, M., Mehra, V., Vrbicky, K., Layle, S., Olson, M.C., Maciel, A., Cullors, A., Garces, J.A., Lukowiak, A.A., 2018. Improved efficacy with targeted pharmacogenetic-guided treatment of patients with depression and anxiety: A randomized clinical trial demonstrating clinical utility. *J Psychiatr Res* 96, 100–107.
- Brown, L.C., Stanton, J.D., Bharthi, K., Maruf, A.A., Muller, D.J., Bousman, C.A., 2022. Pharmacogenomic Testing and Depressive Symptom Remission: A Systematic Review and Meta-Analysis of Prospective, Controlled Clinical Trials. *Clin Pharmacol Ther* 112 (6), 1303–1317.
- Bunka, M., Wong, G., Kim, D., Edwards, L., Austin, J., Doyle-Waters, M.M., Gaedigk, A., Bryan, S., 2023. Evaluating treatment outcomes in pharmacogenomic-guided care for major depression: A rapid review and meta-analysis. *Psychiatry Res* 321, 115102.
- Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315 (7109), 629–634.
- Furukawa, T.A., Cipriani, A., Barbui, C., Brambilla, P., Watanabe, N., 2005. Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol* 20 (1), 49–52.
- Goldberg, J.F., Rosenblatt, J.D., McIntyre, R.S., Preskorn, S.H., de Leon, J., 2019a. Letter to the Editor: Clinical versus statistical significance of pharmacogenomic-guided antidepressant therapy: What's really being measured and marketed? *J Psychiatr Res* 114, 208–209.
- Goldberg, J.F., 2019b. Does Pharmacogenomic Testing Meaningfully Improve Antidepressant Treatment Outcomes When Looking Only at Patients Taking Phase I Hepatically Metabolized Drugs? A Little. *J Clin Psychiatry* 80 (6).
- Green, J.F., Parikh, S.V., Rothschild, A.J., Thase, M.E., Dunlop, B.W., DeBattista, C., Conway, C.R., Forester, B.P., Mondimore, F.M., Shelton, R.C., Macaluso, M., Li, J., Brown, K., Gilbert, A., Burns, L., Jablonski, M.R., Dechairo, B., 2019. Impact of pharmacogenomics on clinical outcomes in major depressive disorder in the GUIDED trial: A large, patient- and rater-blinded, randomized, controlled study. *J Psychiatr Res* 111, 59–67.
- Han, C., Wang, S.M., Bahk, W.M., Lee, S.J., Patkar, A.A., Masand, P.S., Mandelli, L., Pae, C.U., Serretti, A., 2018. A Pharmacogenomic-based Antidepressant Treatment for Patients with Major Depressive Disorder: Results from an 8-week, Randomized, Single-blinded Clinical Trial. *Clin Psychopharmacol Neurosci* 16 (4), 469–480.
- Hieronymus, F., Lisinski, A., Nilsson, S., Eriksson, E., 2019. Influence of baseline severity on the effects of SSRIs in depression: an item-based, patient-level post-hoc analysis. *Lancet Psychiatry* 6 (9), 745–752.
- Higgins, J.P.T., Li, T., Deeks, J.J., 2022. Chapter 6: Choosing effect measures and computing estimates of effect, in: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds), *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). Cochrane, 2022. (Available from www.training.cochrane.org/handbook).
- Jarrett, R.B., Vittengl, J.R., Clark, L.A., 2008. How much cognitive therapy, for which patients, will prevent depressive relapse? *J Affect Disord* 111 (2-3), 185–192.
- Jukic, M., Milosavljevic, F., Molden, E., Ingelman-Sundberg, M., 2022. Pharmacogenomics in treatment of depression and psychosis: an update. *Trends Pharmacol Sci* 43 (12), 1055–1069.
- Kambach, S., Bruelheide, H., Gerstner, K., Gurevitch, J., Beckmann, M., Seppelt, R., 2020. Consequences of multiple imputation of missing standard deviations and sample sizes in meta-analysis. *Ecol Evol* 10 (20), 11699–11712.
- Leucht, S., Fennema, H., Engel, R., Kaspars-Janssen, M., Lepping, P., Szegedi, A., 2013. What does the HAM-D mean? *J Affect Disord* 148 (2-3), 243–248.
- McCarthy, M.J., Chen, Y., Demodena, A., Leckband, S.G., Fischer, E., Golshan, S., Suppes, T., Kelsoe, J.R., 2021. A prospective study to determine the clinical utility of pharmacogenetic testing of veterans with treatment-resistant depression. *J Psychopharmacol* 35 (8), 992–1002.
- Minelli, A., Barlati, S., Baune, B.T., European College of Neuropsychopharmacology, P., Transcriptomics, N., 2022. Evaluating study designs and treatment outcomes of antidepressant pharmacogenetic clinical trials - Challenges and future perspectives. A critical review. *Eur Neuropsychopharmacol* 59, 68–81.
- Oslin, D.W., Lynch, K.G., Shih, M.C., Ingram, E.P., Wray, L.O., Chapman, S.R., Kranzler, H.R., Gelernter, J., Pyne, J.M., Stone, A., DuVall, S.L., Lehmann, L.S., Thase, M.E., Group, P.C.R., Aslam, M., Batki, S.L., Bjork, J.M., Blow, F.C., Brenner, L.A., Chen, P., Desai, S., Dieperink, E.W., Fears, S.C., Fuller, M.A., Goodman, C.S., Graham, D.P., Haas, G.L., Hamner, M.B., Helstrom, A.W., Hurley, R.A., Icardi, M.S., Jurjus, G.J., Kilbourne, A.M., Kreyenbuhl, J., Lache, D.J., Lieske, S.P., Lynch, J.A., Meyer, L.J., Montalvo, C., Muralidhar, S., Ostacher, M.J., Paschall, G.Y., Pfeiffer, P. N., Prieto, S., Przygodzki, R.M., Ranganathan, M., Rodriguez-Suarez, M.M., Roggenkamp, H., Schichman, S.A., Schneeweis, J.S., Simonetti, J.A., Steinhauer, S. R., Suppes, T., Umberto, M.A., Vassy, J.L., Voora, D., Wiechers, I.R., Wood, A.E., 2022. Effect of Pharmacogenomic Testing for Drug-Gene Interactions on Medication Selection and Remission of Symptoms in Major Depressive Disorder: The PRIME Care Randomized Clinical Trial. *JAMA* 328 (2), 151–161.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hrobjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372, n71.
- Papastergiou, J., Quilty, L.C., Li, W., Thiruchselvam, T., Jain, E., Gove, P., Mandlsohn, L., van den Bemt, B., Pojskic, N., 2021. Pharmacogenomics guided versus standard antidepressant treatment in a community pharmacy setting: A randomized controlled trial. *Clin Transl Sci* 14 (4), 1359–1368.
- Perez, V., Salavert, A., Espadaler, J., Tuson, M., Saiz-Ruiz, J., Saez-Navarro, C., Bobes, J., Baca-Garcia, E., Vieta, E., Olivares, J.M., Rodriguez-Jimenez, R., Villagran, J.M., Gascon, J., Canete-Crespillo, J., Sole, M., Saiz, P.A., Ibanez, A., de Diego-Adelino, J., Group, A.-G.C., Menchon, J.M., 2017. Efficacy of prospective pharmacogenetic testing in the treatment of major depressive disorder: results of a randomized, double-blind clinical trial. *BMC Psychiatry* 17 (1), 250.
- Perlis, R.H., Dowd, D., Fava, M., Lencz, T., Krause, D.S., 2020. Randomized, controlled, participant- and rater-blind trial of pharmacogenomic test-guided treatment versus treatment as usual for major depressive disorder. *Depress Anxiety* 37 (9), 834–841.
- Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R., Rushton, L., 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol* 61 (10), 991–996.
- Preskorn, S.H., 2014. Therapeutic Drug Monitoring (TDM) in psychiatry (part I): why studies attempting to correlate drug concentration and antidepressant response don't work. *J Psychiatr Pract* 20 (2), 133–137.
- Samara, M.T., Dold, M., Gianatsi, M., Nikolakopoulou, A., Helfer, B., Salanti, G., Leucht, S., 2016. Efficacy, Acceptability, and Tolerability of Antipsychotics in Treatment-Resistant Schizophrenia: A Network Meta-analysis. *JAMA Psychiatry* 73 (3), 199–210.
- Shan, X., Zhao, W., Qiu, Y., Wu, H., Chen, J., Fang, Y., Guo, W., Li, L., 2019. Preliminary Clinical Investigation of Combinatorial Pharmacogenomic Testing for the Optimized Treatment of Depression: A Randomized Single-Blind Study. *Front Neurosci* 13, 960.
- Singh, A.B., 2015. Improved Antidepressant Remission in Major Depression via a Pharmacokinetic Pathway Polygene Pharmacogenetic Report. *Clin Psychopharmacol Neurosci* 13 (2), 150–156.
- Sterne, J.A.C., Savovic, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., Cates, C. J., Cheng, H.Y., Corbett, M.S., Eldridge, S.M., Emberson, J.R., Hernan, M.A., Hopewell, S., Hrobjartsson, A., Junqueira, D.R., Juni, P., Kirkham, J.J., Lasserson, T., Li, T., McAleenan, A., Reeves, B.C., Shepperd, S., Shrier, I., Stewart, L.A., Tilling, K., White, I.R., Whiting, P.F., Higgins, J.P.T., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 366, 14898.
- Stuart, A., Ord, J.K., 1998. *Kendall's Advanced Theory of Statistics*, sixth Ed, 1. Arnold, London.
- Sun, Y., Fu, Z., Bo, Q., Mao, Z., Ma, X., Wang, C., 2020. The reliability and validity of PHQ-9 in patients with major depressive disorder in psychiatric hospital. *BMC Psychiatry* 20 (1), 474.
- Tiwari, A.K., Zai, C.C., Altar, C.A., Tanner, J.A., Davies, P.E., Traxler, P., Li, J., Cogan, E. S., Kucera, M.T., Gugila, A., Braganza, N., Emmerson, H., Zai, G., Muller, D.J., Levitan, R., Kloiber, S., Daskalakis, Z.J., Frey, B.N., Bowen, J.M., Tarride, J.E., Tytus, R., Chandrasena, R., Voudouris, N., Taylor, V.H., Tempier, R., Sharma, V., Vasudev, A., Dzongowski, P., Pliamm, L., Greenspoon, T., Dechairo, B.M., Kennedy, J.L., 2022. Clinical utility of combinatorial pharmacogenomic testing in depression: A Canadian patient- and rater-blinded, randomized, controlled trial. *Transl Psychiatry* 12 (1), 101.
- Tonk, E.C.M., Gurwitz, D., Maitland-van der Zee, A.H., A, C., Janssens, J.W., 2017. Assessment of pharmacogenetic tests: presenting measures of clinical validity and potential population impact in association studies. *Pharmacogenomics J* 17, 386–392.
- van der Schans, J., Hak, E., Postma, M., Breuning, L., Brouwers, J., Ditters, K., Jansen, P. A.F., Kok, R.M., Maring, J.G., van Marum, R., Mulder, H., Nanninga, J., Oude Voshaar, R.C., Risselada, A.J., Vleugel, L., Stek, M., van Schaik, R.H.N., Berm, E.J.J., Wilffert, B., 2019. Effects of Pharmacogenetic Screening for CYP2D6 Among Elderly Starting Therapy With Nortriptyline or Venlafaxine: A Pragmatic Randomized Controlled Trial (CYSCE Trial). *J Clin Psychopharmacol* 39 (6), 583–590.
- Vande Voort, J.L., Orth, S.S., Shekunov, J., Romanowicz, M., Geske, J.R., Ward, J.A., Leibman, N.I., Frye, M.A., Croarkin, P.E., 2022. A Randomized Controlled Trial of Combinatorial Pharmacogenetics Testing in Adolescent Depression. *J Am Acad Child Adolesc Psychiatry* 61 (1), 46–55.
- Vos, C.F., Ter Hark, S.E., Schellekens, A.F.A., Spijker, J., van der Meij, A., Grotenhuis, A. J., Mihaescu, R., Kievit, W., Donders, R., Aarnoutse, R.E., Coenen, M.J.H., Janzing, J. G.E., 2023. Effectiveness of Genotype-Specific Tricyclic Antidepressant Dosing in Patients With Major Depressive Disorder: A Randomized Clinical Trial. *JAMA Netw Open* 6, e2312443.
- Wang, X., Wang, C., Zhang, Y., An, Z., 2023. Effect of pharmacogenomics testing guiding on clinical outcomes in major depressive disorder: a systematic review and meta-analysis of RCT. *BMC Psychiatry* 23 (1), 334.
- Weir, C.J., Butcher, I., Assi, V., Lewis, S.C., Murray, G.D., Langhorne, P., Brady, M.C., 2018. Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review. *BMC Med Res Methodol* 18, 25.
- Winner, J.G., Carhart, J.M., Altar, C.A., Allen, J.D., Dechairo, B.M., 2013. A prospective, randomized, double-blind study assessing the clinical impact of integrated pharmacogenomic testing for major depressive disorder. *Discov Med* 16 (89), 219–227.
- Zeier, Z., Carpenter, L.L., Kalin, N.H., Rodriguez, C.I., McDonald, W.M., Widge, A.S., Nemeroff, C.B., 2018. Clinical Implementation of Pharmacogenetic Decision Support Tools for Antidepressant Drug Prescribing. *Am J Psychiatry* 175 (9), 873–886.
- Zubenko, G.S., Sommer, B.R., Cohen, B.M., 2018. On the Marketing and Use of Pharmacogenetic Tests for Psychiatric Treatment. *JAMA Psychiatry* 75 (8), 769–770.