3-2023

Evaluating treatment outcomes in pharmacogenomic-guided care for major depression: A rapid review and meta-analysis.

Mary Bunka

Gavin Wong

Dan Kim

Louisa Edwards

Jehannine Austin

See next page for additional authors

Recommended Citation

Bunka M, Wong G, Kim D, et al. Evaluating treatment outcomes in pharmacogenomic-guided care for major depression: A rapid review and meta-analysis. Psychiatry Res. 2023;321:115102. doi:10.1016/j.psychres.2023.115102

Psychiatry Research 321 (2023) 115102

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Review article

Evaluating treatment outcomes in pharmacogenomic-guided care for major depression: A rapid review and meta-analysis

Mary Bunka [a,b], Gavin Wong [a,b,*], Dan Kim [a,b], Louisa Edwards [a,b], Jehannine Austin [c], Mary M. Doyle-Waters [b], Andrea Gaedigk [d,e], Stirling Bryan [a,b]

[a] School of Population and Public Health, University of British Columbia (UBC), 828 West 10th Avenue, Research Pavilion, 7th Floor, Vancouver, BC V5Z 1M9, Canada
[b] Centre for Clinical Epidemiology and Evaluation, Vancouver Coastal Health Research Institute, Vancouver, BC, Canada
[c] BC Mental Health and Substance Use Services Research Institute, UBC, Vancouver, BC, Canada
[d] Division of Clinical Pharmacology, Toxicology & Therapeutic Innovation, Children's Mercy Kansas City, Kansas City, MO, USA
[e] Department of Pediatrics, University of Missouri-Kansas City, Kansas City, MO, USA

ARTICLE INFO

Keywords:
Systematic review
Depression
Mental health
Antidepressant
Pharmacotherapy
Pharmacogenetics
Dose-response relationship
Drug

ABSTRACT

Pharmacogenomic (PGx) testing may increase the probability of remission and response in patients with major depressive disorder (MDD) undergoing pharmacotherapy. Given the potential implications of these outcomes and recent proliferation of PGx studies, we conducted a systematic review to evaluate the effectiveness of PGx testing on clinical outcomes in patients with MDD as compared to treatment as usual (TAU). MEDLINE, Embase, PsycInfo, and CENTRAL were searched for English-language articles from 2000 to 2021 for randomized controlled trials (RCTs) comparing PGx-guided treatment vs. TAU in patients with MDD. Meta-analyses were conducted in R. Ten RCTs were included: eight reported remission and seven reported response. The best available evidence suggests that PGx-guided care for moderate-to-severe adult depression is more likely to result in remission and response than TAU (both risk ratios significant). However, there are limitations in the evidence base, including high risk of bias and inconsistency between trials. Despite the consequent very low certainty in the magnitude of effect, there is confidence in the direction. Though modest, the beneficial effects of PGx for adults with moderate-severe MDD could – as a result of the scope and scale of the condition and its impacts – have important ramifications for patients and the health system.

1. Introduction

Major Depressive Disorder (MDD) is a commonly occurring mental health condition that seriously impacts functioning across many aspects of life and has a high recurrence rate. Globally, MDD is among the leading causes of disability (World Health Organization, 2017). In Canada, the lifetime prevalence of MDD is 11.2% (Knoll and MacLennan, 2017). Although effective pharmacological treatment is available, fewer than half of patients with MDD respond to the first medication they are prescribed (Ruhe et al., 2006), and more than 30% of patients still experience depression symptoms after trying several medications (Rush et al., 2006a). This can lead to a period of trial-and-error prescribing, which results in poorer long-term outcomes for patients who undergo several trials (Rush et al., 2006a) and is an additional financial burden to the health system. Thus, interventions that bring about remission or response, which lead to fewer medication trials are likely to be beneficial on both the individual patient and healthcare system levels.

Genetic variation can influence how medications, including antidepressants, are metabolized. Studies have suggested that up to 42% of the variation in treatment effect can be attributed to genetic differences (Rush et al., 2006b; Tansey et al., 2013). Pharmacogenomic (PGx) testing, therefore, holds great theoretical appeal as a means of improving pharmacological treatment. PGx tests aim to detect variants in genes involved in drug metabolism and response, that can inform recommendations based on variants found in a patient to guide treatment, an approach also known as precision medicine or individualized drug therapy, to maximize therapeutic outcome. The Dutch Pharmacogenetics Working Group provided prescribing guidelines and has published two papers with recommendations (Swen et al., 2008, 2011).

* Corresponding author at: School of Population and Public Health, University of British Columbia (UBC), 828 West 10th Avenue, Research Pavilion, 7th Floor, Vancouver, BC V5Z 1M9, Canada.
E-mail address: gavin.wong@ubc.ca (G. Wong).

https://doi.org/10.1016/j.psychres.2023.115102
Received 7 November 2022; Received in revised form 13 January 2023; Accepted 6 February 2023
Available online 8 February 2023

Updates to these guidelines are posted on the Royal Dutch Pharmacists Association (KNMP) website. The Clinical Pharmacogenetics Implementation Consortium (CPIC) has also released guidelines for antidepressants, including selective serotonin reuptake inhibitors and tricyclic antidepressants, based on evidence that variation in the *CYP2C19* and *CYP2D6* pharmacogenes – the two genes evaluated by all PGx decision-support tools for antidepressants (Bousman and Hopwood, 2016; Bousman et al., 2017) – contribute to variable drug metabolism (Hicks et al., 2015, 2017). CPIC also provides a gene-drug table, in which one can look up antidepressants or other drugs of interest, the PharmGKB level of evidence, and related PGx-based prescribing recommendations for that drug (Clinical Pharmacogenetics Implementation Consortium, 2022).

Given the significant potential benefits of PGx testing on the probability of remission (a score below clinical cut-off for diagnosis on a validated measure of depression), response (a 50% decrease in scale score on a validated measure of depression), and medication tolerability for those with MDD, several clinical trials (Winner et al., 2013; Singh, 2015; Perez et al., 2017; Bradley et al., 2018; Han et al., 2018; Greden et al., 2019; Shan et al., 2019; van der Schans et al., 2019; Perlis et al., 2020; Ruano et al., 2020; Vande Voort et al., 2021) have been conducted to evaluate their effectiveness. The individual randomized control trial (RCT) results are mixed. Recent reviews have found positive associations between PGx testing and remission and response compared to treatment as usual (Rosenblat et al., 2017, 2018; Bousman et al., 2019; Brown et al., 2020), though meta-analysis was not always conducted and reviews were limited by non-systematic searches, non-replicable methodologies, or by industry affiliations. Therefore, an overall pooled assessment of the evidence for the use of PGx testing in MDD care is needed.

A recent health technology assessment conducted in Canada recommended against implementing PGx for depression (Ontario Health, 2021). However, this assessment divided the PGx tests according to the supplier, which may have reduced the power to detect a difference between pharmacogenomic testing in general and treatment as usual owing to differences in the variants tested among the suppliers.

PGx tests themselves provide information that aids clinicians in making a prescription choice. By combining PGx tests from different suppliers in a meta-analysis, we are assessing the efficacy of "guided treatment". This is the gap the current paper is designed to address. Our goal is to examine, for MDD patients, whether or not PGx tests as a whole are effective compared to treatment as usual by examining clinical outcomes in patients eligible for pharmacotherapy who have undergone PGx testing versus treatment as usual. Our review adds to the knowledge base by including the most up-to-date RCTs and running meta-analyses and critical appraisals of all included outcomes.

2. Methods

This is a rapid review that follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Guidelines (Page et al., 2021). The protocol was registered on PROSPERO (#204827).

Eligible studies included patients aged 6 and older with MDD who were eligible for treatment with pharmacotherapy. Randomized controlled trials (RCTs) published from 2000 to 2021 which included a comparison of PGx testing (any combinatorial or single-gene test designed to help select the optimal drug with which to treat a patient) with treatment as usual were included. There were no restrictions on comorbidities as long as depression was the primary disorder undergoing treatment. Exclusion criteria included non-English language publications; commentaries, editorials, conference abstracts/proceedings, other non-peer-reviewed articles; studies without an appropriate comparator; studies which involved PGx testing for physical conditions; and studies that included unspecified PGx testing (that is, testing for any/all possible conditions, history, or any/all medications). Studies of perinatal depression were also excluded.

The search strategy (available in the online supplement) was developed by an information specialist with inputs from reviewers. Published articles were identified in MEDLINE, Embase, Register of Controlled Trials (CENTRAL) *via* Ovid and APA PsycInfo through EBSCO. Additionally, reference lists of included RCTs were hand-searched to identify any missing eligible papers. The searches were run sequentially between July and October 2020. An updated search for recent publications was repeated July 26, 2021. One reviewer (MB) screened all of the abstracts, while a second reviewer (DK) checked 8% of the abstracts for agreement. These 8% were identified during the first screening as more uncertain than others and worthy of a second look. The first and second reviewers read each of the full texts and independently marked them for inclusion or exclusion. Disagreements were resolved through discussion until consensus was reached. The first reviewer extracted data from the RCTs and the second reviewer cross-checked for agreement. Any discrepancies were resolved through discussion. A standard data extraction sheet was used. The number of events or percentage was extracted for response (defined as a $\geq$50% decrease in Hamilton Depression Scale-17 Item Version (Hamilton, 1960) [HAM-D17] score) (Winner et al., 2013; Perez et al., 2017; Bradley et al., 2018; Greden et al., 2019; Shan et al., 2019; Han et al., 2020; Perlis et al., 2020) and remission (defined as a score of $\leq$7 on the HAM-D17) (Winner et al., 2013; Singh, 2015; Perez et al., 2017; Bradley et al., 2018; Greden et al., 2019; Shan et al., 2019; Han et al., 2020; Perlis et al., 2020), as well as the number of treatment changes, total adverse effects, serious adverse events, total discontinuation from the study, withdrawal due to adverse effects, and mortality. We chose to compare the number of people who remitted and responded on the HAM-D17 scale as it was the most widely reported between the included RCTs. Means and standard deviations were obtained for continuous outcomes, including change in depression scale score, adherence to medication, and quality of life. Baseline characteristics including sex, age, type of PGx test, and study characteristics were also extracted.

The risk of bias (RoB) in the included RCTs was assessed using the Cochrane Risk of Bias Tool (Version 2; RoB 2) (Sterne et al., 2019). Two reviewers independently assessed each outcome in each RCT and then checked agreement, resolving disagreements through discussion.

Dichotomous outcomes were analyzed using risk ratio (RR). We calculated risk difference and number needed to treat for each statistically significant RR when possible. Continuous outcomes, such as mean change in depression scale score, were not meta-analyzed due to a lack of reported data.

RCTs that provided data for the number of patients experiencing remission and response were included in the meta-analyses (Winner et al., 2013; Singh, 2015; Perez et al., 2017; Bradley et al., 2018; Han et al., 2018; Greden et al., 2019; Shan et al., 2019; Perlis et al., 2020). All analyses were based on the intention-to-treat principle. Patients who discontinued the randomized treatment were considered to be non-responders and non-remitters at the end of the study. Random effects meta-analysis was conducted for response and remission outcomes using the "meta" package in R version 4.0.5.; results with 95% confidence intervals and p-values were reported. Fixed effect meta-analyses were conducted for other outcomes with less heterogeneity. Statistical heterogeneity was assessed with I-square. We also calculated risk difference and number needed to treat for each statistically significant outcome when appropriate.

Certainty in the evidence was assessed using the GRADEpro Guideline Development Tool (GDT) (Evidence Prime Inc., 2020). Risk of bias due to missing results from reporting bias was explicitly addressed in this tool.

3. Results

The search returned 2485 abstracts for screening (Fig. 1). An additional 24 abstracts were identified through hand-searching. Ten RCTs published in 14 articles (Winner et al., 2013; Singh, 2015; Perez et al.,

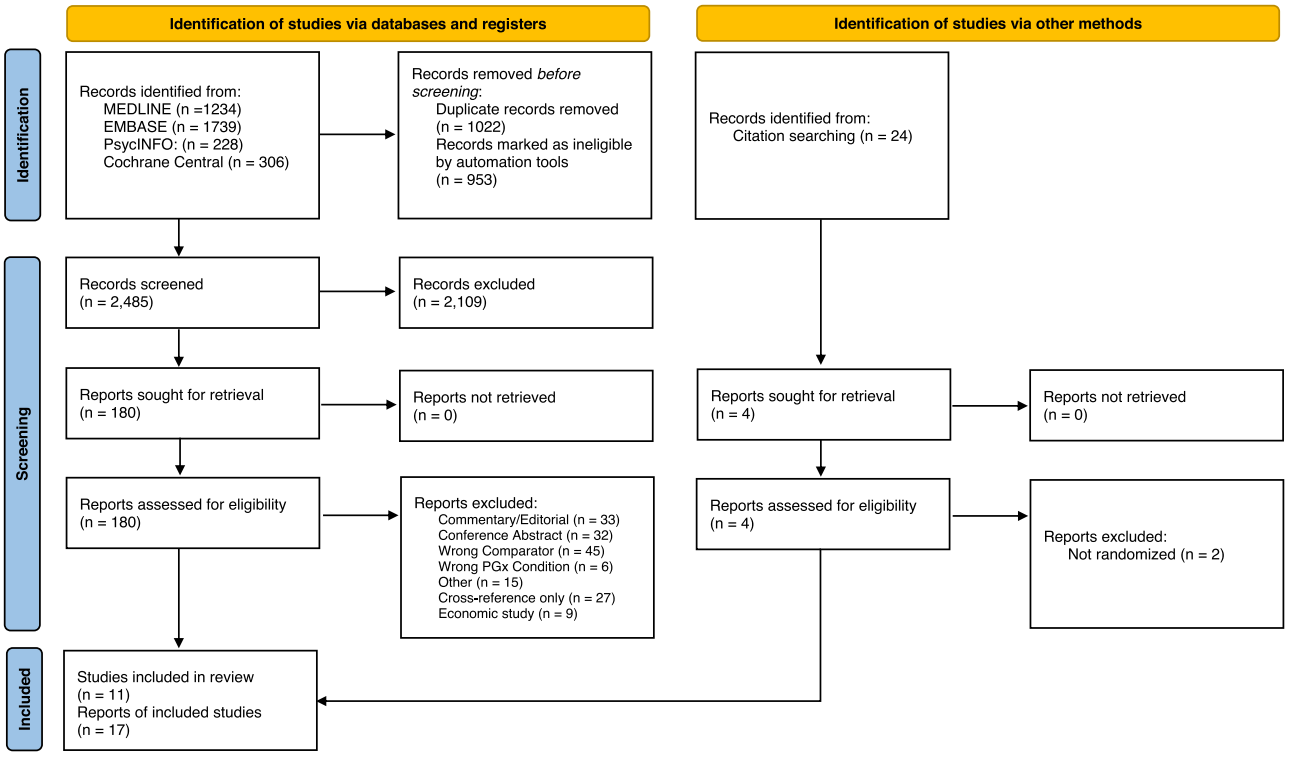

Fig. 1. PRISMA diagram of study selection.

2017; Bradley et al., 2018; Han et al., 2018; Dunlop et al., 2019; Greden et al., 2019; Menchon et al., 2019; Shan et al., 2019; Thase et al., 2019; van der Schans et al., 2019; Forester et al., 2020; Perlis et al., 2020; Ruano et al., 2020) were included in the review. For RCTs associated with more than one published paper, data were extracted and meta-analyzed from the primary publication. After running an updated search in July 2021, one additional RCT (Vande Voort et al., 2021) was included. However, it was not meta-analyzed with the other RCTs as it included an adolescent-only population, while the others enrolled adults only. One additional paper (Ruano et al., 2021), a subgroup analysis of the CYP-GUIDES trial (Ruano et al., 2020), was also included as was a corrigendum (Han et al., 2020) to one of the previously included RCTs (Han et al., 2018). Excluded studies and reasons for their exclusion can be found in the supplementary material.

Ten RCTs with adult participants (n=4333) met the inclusion criteria (Winner et al., 2013; Singh, 2015; Perez et al., 2017; Bradley et al., 2018; Han et al., 2018; Greden et al., 2019; Shan et al., 2019; van der Schans et al., 2019; Perlis et al., 2020; Ruano et al., 2020). Each RCT included a PGx testing arm and an active control arm (treatment as usual), with follow-up periods ranging from 8 to 12 weeks. The majority of the included RCTs were double-blinded for the two main outcomes, response and remission; a blinded assessor took HAM-D17 scores, instead of the prescriber/caregiver, who could not be blinded to treatment condition given the nature of the intervention. Two RCTs did not utilize a blinded reviewer for the response and remission outcomes presented here (Perez et al., 2017; Han et al., 2020). Eight RCTs reported on remission as an outcome and seven reported on response. Only one RCT reported on patient quality of life (van der Schans et al., 2019). A comprehensive description of trial designs of each RCT can be found in Supplementary Table 1.

Baseline characteristics were similar between arms, with one exception: more females were allocated to the control group versus the intervention group in Winner et al. (2013): 92% and 69%, respectively. Mean baseline depression score on the HAM-D17 ranged from 19.47 to 24.81 in the intervention groups and 19.01 to 24.66 in the control groups. Every included RCT had a larger female population than male, the lowest proportion of female participants being 51.3% and 50.3% in the intervention and control groups, respectively (Ruano et al., 2020). Most participants were middle-aged. Detailed information on baseline characteristics by RCT are summarized in Supplementary Table 2. The reporting of other baseline characteristics which might influence the outcomes, especially ethnicity, was sporadic. It is worth noting that no RCT reported on gender, a social construction of expression and identity, as opposed to sex, which is a biological descriptor and was included in each RCT. Only one RCT reported average length of the current depressive episode and mean number of previous episodes (Singh, 2015), one RCT reported baseline body mass index (BMI) (Perlis et al., 2020), and one RCT reported marital status and religion (Han et al., 2018). Two RCTs (Singh, 2015; Han et al., 2018) reported on employment status.

In a meta-analysis of the eight RCTs (Winner et al., 2013; Singh, 2015; Perez et al., 2017; Bradley et al., 2018; Han et al., 2018; Greden et al., 2019; Shan et al., 2019; Perlis et al., 2020) (total n= 2341) that provided data on number of patients with depression remission at the end of study period (defined as a non-clinical HAM-D17 scale score of <7), the RR of remission was 1.46 (95% CI: 1.02-2.08, p=0.043) in the intervention (PGx) arm compared with treatment as usual (TAU) control arm (Fig. 2). The absolute risk difference (RD) was 0.124 and the Number Needed to Treat (NNT) was 8. Seven RCTs (Winner et al., 2013; Perez et al., 2017; Bradley et al., 2018; Han et al., 2018; Greden et al., 2019; Shan et al., 2019; Perlis et al., 2020) (total n= 2188) reported on response (a $\geq$ 50% reduction in HAM-D17 scale score from baseline to after treatment). RR of response was 1.32 (95% CI: 1.00-1.73, p=0.047) in the PGx group versus the TAU group (Fig. 3). The RD was 0.126 with a NNT of 8. To test these results, we performed subgroup analyses of response and remission with the high RoB RCTs removed. This analysis revealed that, as the sample size decreased, uncertainty increased until the results were no longer significant. However, the point estimates were similar between analyses (see Supplemental Figs. 1 and 2).

Total study discontinuation, analyzed with a fixed effect model (Fig. 4), was slightly lower in the PGx group, though the result was not statistically significant with a RR of 0.89 (95% CI: 0.78-1.01, p=0.07).

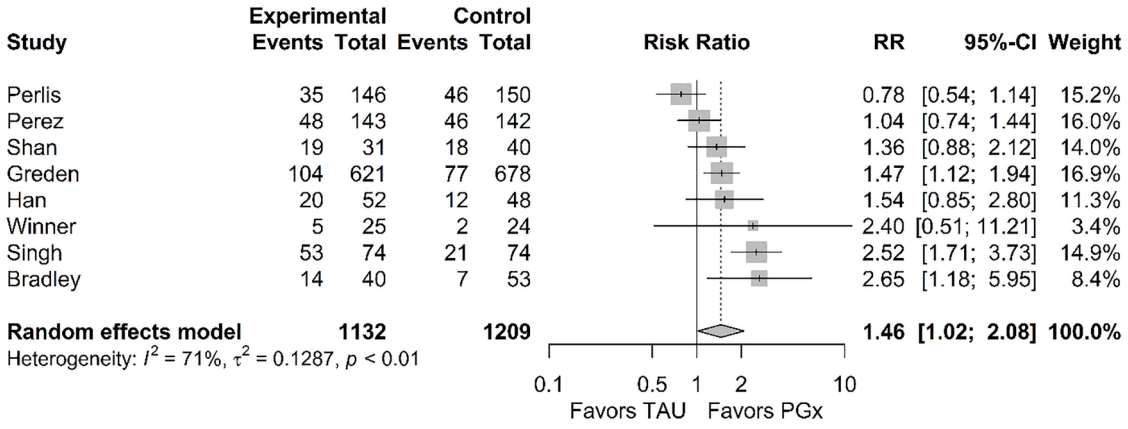

Fig. 2. Forest plot showing remission meta-analysis and risk ratio.
Abbreviations: RR: risk ratio, CI: confidence interval, TAU: treatment as usual (control arm), PGx: pharmacogenomic testing (intervention arm).

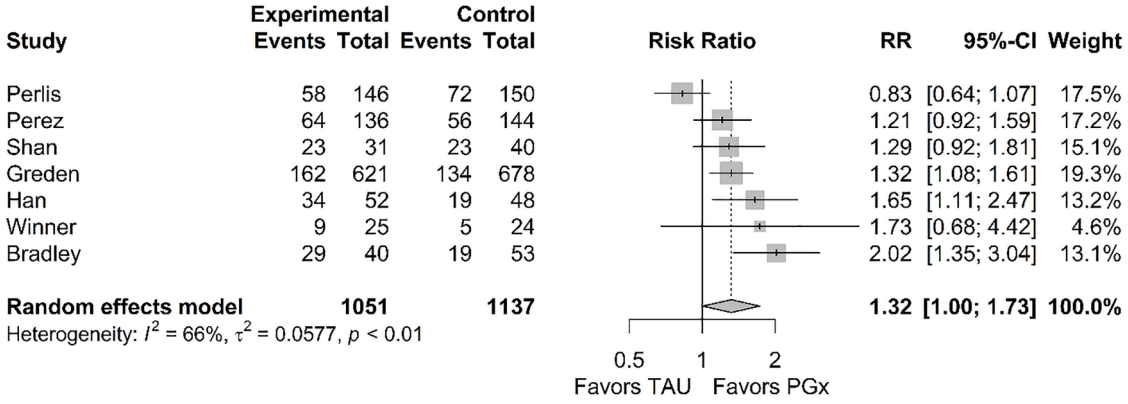

Fig. 3. Forest plot showing response meta-analysis and risk ratio.
Abbreviations: RR: risk ratio, CI: confidence interval, TAU: treatment as usual (control arm), PGx: pharmacogenomic testing (intervention arm).

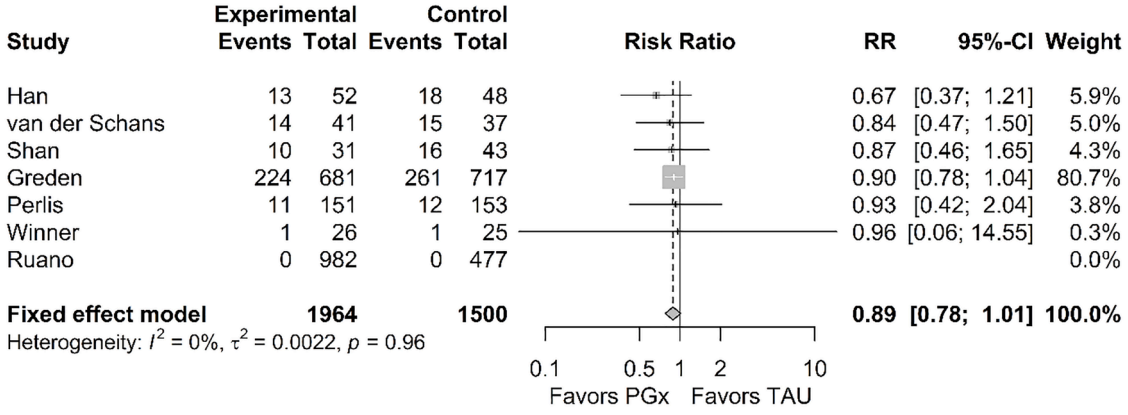

Fig. 4. Forest plot showing total study discontinuation meta-analysis and risk ratio.
Abbreviations: RR: risk ratio, CI: confidence interval, TAU: treatment as usual (control arm), PGx: pharmacogenomic testing (intervention arm).

Serious adverse effects and withdrawal due to adverse effects were also not significant, though only two RCTs reported data for these outcomes.

Risk of bias was assessed by individual outcome according to RoB 2. For both remission and response, four (Perez et al., 2017; Bradley et al., 2018; Han et al., 2018; Shan et al., 2019) included RCTs had a high risk of bias, largely due to missing outcome data and risk of bias in the outcome measurement due to unblinded assessors. It is worth noting that in the RoB 2 tool, each outcome is assessed for risk of bias separately, meaning that RCTs could include outcomes with different RoB ratings (see Supplemental Figs. 3–5 for the RoB 2 traffic plots, created with the "robvis" tool for R) (McGuinness and Higgins, 2021). Because of the overall high risk of bias (in four out of eight RCTs meta-analyzed for remission and four out of seven meta-analyzed for response), and inconsistency (which, in the context of the GRADEpro tool, refers to wide confidence intervals and unexplained heterogeneity) we arrived at a very low certainty of evidence as assessed by GRADEpro (Table 1) (Evidence Prime Inc., 2020). Therefore, the point estimates presented above should be interpreted with caution. Although the evidence points

Table 1
GRADE summary of evidence.

Certainty assessment							Probability of outcome	Effect		Certainty	Importance	
No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Standard of care	Relative (95% CI)	Absolute (95% CI)		
Remission (assessed with: HAM-D17)												
8	randomized trials	serious[a]	very serious[b]	not serious	not serious	none[c]	26.2%	RR 1.46 (1.02 to 2.08)	121 more per 1000 (from 5 more to 283 more)	⨁◯◯◯ Very low	Critical	
Response (assessed with: HAM-D17)												
7	randomized trials	serious[d]	very serious[e]	not serious	not serious	none[c]	41.4%	RR 1.32 (1.00 to 1.73)	132 more per 1000 (from 0 fewer to 302 more)	⨁◯◯◯ Very low	Critical	
Total Study Discontinuation												
6	randomized trials	not serious	not serious	serious[f]	not serious	none[c]	25.8%	RR 0.89 (0.78 to 1.01)	28 fewer per 1000 (from 57 fewer to 3 more)	⨁⨁⨁◯ Moderate	Important	

Explanations:
[a] Three of the top four studies by weight had low or only "some concerns." The second largest study had a high risk of bias. Overall, missing outcome data (due to unexplained participant loss) and risk of bias due to measurement (single-blinding and/or patient-reported outcomes) warrant rating the evidence down by one level.
[b] High heterogeneity between studies ($I^2 = 71\%$) that was not accounted for after removing "obvious" outliers.
[c] Not enough studies for analysis of publication bias; funnel plots not created for this reason.
[d] Crucial limitation for one criterion, or some limitations for multiple criteria, sufficient to lower confidence in the estimate of effect. (from GRADE handbook).
[e] . Moderate to high heterogeneity between studies ($I^2 = 66\%$, $p < 0.01$). Heterogeneity is 11% when Perlis et al. study is removed from the meta-analysis. Singh study not included in this analysis; difference in heterogeneity between remission and response might be due to this missing study and not due to a more homogenous sample.
Abbreviations: No.: number, CI: confidence interval, HAM-D17: Hamilton Depression Rating Scale, 17 Item Version, RR: risk ratio.

to a positive relationship between PGx-guided treatment and response/remission, the magnitude of this effect may change as more evidence becomes available. The meta-analysis of total discontinuation included seven RCTs, two of which had a low risk of bias rating and two of which had a high risk of bias rating. All data were accounted for, so each RCT received a "low" rating for risk of bias due to missing data. The two RCTs with high risk of bias for this outcome were rated this way because of knowledge of forthcoming randomization allocations and because of deviations from the intended intervention (van der Schans et al., 2019; Ruano et al., 2020).

The only RCT including adolescents (Vande Voort et al., 2021) found no significant differences between groups in symptom improvement (that is, change in depression scale score), response, or remission at week 8 or at any point throughout the study as measured with the Children's Depression Rating Scale-Revised (CDRS-R) or the Quick Inventory of Depressive Symptomatology (QIDS). There was no statistically significant difference in the number of adverse events or side effects between groups. While there was a statistically significant improvement in patient and parent satisfaction with care in the overall study population, it was not significantly different between treatment arms. Based on these findings, there is currently no evidence to support the use of PGx tests in depression care for adolescents.

4. Discussion

Given the recent proliferation in RCTs evaluating the effectiveness of using PGx to guide pharmacological treatment, this review provides a timely synthesis of the best available evidence. The goal of antidepressant treatment is to bring about remission in depression. In the absence of remission, a response to medication can be an indicator of some symptom relief that resulted in a lower depression scale score.

Our results suggest that use of PGx to guide antidepressant treatment could increase the rate of remission in adults with moderate to severe depression, whereby 121 additional people remit per 1000 treated when compared to treatment without PGx testing. Similarly, 132 additional patients could respond per 1000 if treatment is informed by PGx compared to treatment as usual without PGx guidance. The importance of this finding, that more patients could remit and respond to treatment guided by PGx, cannot be overstated. A patient's quality of life increases when in remission compared to when in a depressive episode (Bansback et al., 2012; IsHak et al., 2015; Kolovos et al., 2017). Implications of improvement in quality of life reach beyond the individual patient level and into the systems level as more patients reaching remission will likely be associated with less overall use of the healthcare system as well as decreased absenteeism. Improving the likelihood of response and remission from major depression is undoubtedly a positive outcome and, in this regard, the evidence points to a benefit of PGx-guided treatment over standard of care.

We do acknowledge limitations in the evidence, such as the possibility of a type 1 error. Certainty in the evidence was assessed as very low for both response and remission due to high risk of bias and inconsistency in the included RCTs. We conducted preliminary explorations of both of these limitations in the evidence. As mentioned above, we performed subgroup analyses in response and remission with the high risk of bias RCTs removed. Importantly, the point estimate remained consistent for response and remission when only RCTs with risk of bias ratings of "low" and "some concerns" were included in the meta-analyses. Of course, with the decrease in sample size, this analysis also decreased certainty in the outcomes, as is expected (Supplemental Figs. 1 and 2). Given that removing high risk of bias-rated RCTs resulted in only four RCTs in the remission meta-analysis and three in the response meta-analysis, it is clear that additional high-quality trials are warranted.

We also examined inconsistency by identifying outliers in the response and remission meta-analyses. Each meta-analysis was repeated without the identified outliers and heterogeneity dropped from $I^2 = 71\%$ to $I^2 = 20\%$ in remission and from $I^2 = 66\%$ to $I^2 = 11\%$ in response. In the absence of sufficient RCTs to conduct a network meta-analysis or create a funnel plot, observing the remaining studies on a forest plot (Supplemental Figs. 6 and 7) suggests that a single population mean is emerging, instead of multiple means, as might be expected if each test was not comparable. While this is not proof of a class effect, it is

preliminary evidence supporting one single population mean (i.e., a class effect). Again, additional high-quality RCTs and other analyses are needed to corroborate our findings.

Our review presents the most up-to-date evidence, including the first RCT of PGx for depression care in adolescents. Furthermore, our review follows the new Cochrane Risk of Bias assessment guidelines, appraising each outcome instead of each RCT included in the review. Our findings are in line with other recent reviews that conducted meta-analyses of remission and response, and reported that PGx-guided treatment was associated with increased remission (RRs ranging from 1.41 to 1.74) (Rosenblat et al., 2018; Brown et al., 2020, 2022) and response (RRs ranging from 1.36 to 1.40) (Rosenblat et al., 2018; Brown et al., 2020). This strongly suggests that there is a positive effect of PGx testing on remission and response in MDD, despite the limitations in existing data. Some of these reviews did not include meta-analyses (Rosenblat et al., 2017) or analyzed each supplier's PGx test separately and did not find significant outcomes (Ontario Health, 2021). By making a reasonable assumption about assessing PGx testing as a "class" of treatment guidance, we found significant effects with important implications for patients. However, Brown et al. (2022) carried out a similar review and meta-analysis to ours, with similar findings. Yet, there are important differences between these two studies. First, Brown et al. (2022) only included remission in their meta-analysis, while we evaluated a wider range of clinically important outcomes, including remission, response to treatment, withdrawal due to adverse effects, total discontinuation, and serious adverse events. This collection of outcomes provides a more comprehensive view of the evidence on this topic. Secondly, we included patients under the age of 18 as our target population, which represents a broader clinical population than Brown et al. (2022). As such, our review contains the only adolescent RCT of pharmacogenomics in MDD that has been published to date and, although only one RCT was identified, this finding fills an important gap in the evidence base.

Though our review presents the best available published evidence, there are several limitations to consider in addition to the points raised above regarding risk of bias and heterogeneity. Evidence was only available for adult patients (with the exception of one recent RCT (Vande Voort et al., 2021)), specifically those with moderate to severe depression. The RCTs included ethnically homogenous populations and usually included substantially more females than males in their analyses. Future research should prioritize diverse populations that are more representative of actual populations with MDD, including patients with mild MDD, children, and adolescents. Future research should also capture important demographic information such as gender and ethnicity. In addition, inconsistency in reporting outcome data as specified in study protocols, as well as inconsistencies between RCTs in how data were reported (i.e., reporting data with the same measure) limited the number of meta-analyses we were able to perform. For example, although eight RCTs reported scores on the HAM-D17, only two presented the data in the same way or with enough detail to conduct a meta-analysis. RCTs reported some combination of mean baseline score with mean endpoint score, mean decrease in scale score, or mean percentage decrease in scale score. Understanding how PGx testing affects symptom scores is paramount, as remission is not always possible and a change in depression scale score could represent meaningful symptom reduction for patients. Standardizing conventions for reporting this information would increase robustness of future meta-analyses. Risk of bias due to outcome measurement was also concerning; while remission and response are dichotomous outcomes, the tool used to measure them, the HAM-D17, is a clinician-rated scale that involves some amount of judgment. The use of unblinded assessors, especially in industry-sponsored (or industry-supported) trials may have influenced the results for response and remission. Future investigations need to be conducted with attention paid to substantially reducing the amount of bias in the trials.

In addition to examining the effectiveness of PGx across a broader group of participants, future research should address the current lack of evidence around adverse effects and serious adverse events, as these were not sufficiently reported to allow for meta-analysis. Avoiding side effects is a high priority for patients and one of the most common reasons patients choose to discontinue a medication prematurely (Ho et al., 2017). Also, the lack of evidence should be addressed by recording and reporting all adverse events in both study arms, regardless of significance, as a pooled effect may be found between studies. Finally, as the longest blinded follow-up period was 12 weeks, future trials need to gage the long-term efficacy of PGx testing for depression, including rates of recurrence. Depression is a recurring disorder; patients who reach remission could experience another depressive episode later in life (Rush et al., 2006b; Knoll and MacLennan, 2017; World Health Organization, 2017). A long-term study could provide much-needed evidence on the effects of PGx-guided treatment on relapse as well as time to recurrence, both with massive implications for patients and the healthcare system.

The evidence broadly supports PGx-guided treatment for depression in terms of response and remission outcomes. Discontinuing study participation was found to be slightly less likely with PGx-guided care which, if viewed as a proxy for treatment adherence and satisfaction with care, points favorably towards PGx-guided treatment. Overall, it is clear that additional high-quality research is warranted, with trialists considering ways to minimize participant attrition and potential risk of detection bias, and moving to public or independent funding sources to minimize conflict of interest. This is a dynamic clinical area, in which the evidence is quickly evolving and new technologies are being developed and strengthened while concurrently being used to guide treatment decisions (Oslin et al., 2022; Scherf-Clavel et al., 2022; Tiwari et al., 2022). High quality evidence that measures long term patient-oriented outcomes can help fill the current knowledge gap and improve patient care in the future.

Author contributions

All authors meet the ICMJE criteria for authorship. LE, JA, MDW, AG, and SB designed the research question. MDW conducted the literature search. MB and DK acted as the first and second reviewers, respectively, and conducted a critical appraisal. MB conducted meta-analyses with supervision by GW. MB drafted the manuscript, which GW, DK, LE, JA, MDW, AG, and SB revised and/or approved of. All authors have reviewed and approved the final manuscript for publication.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the larger PGx4Dep team for their contributions to the research design and manuscript preparation and revisions, with special thanks to Lisa Ridgway (patient partner), Alison Hoens, and Kim McGrail. The authors offer gratitude to the Coast Salish Peoples, including the xʷməθkwəy̓əm (Musqueam), Skwxwú7mesh (Squamish), and Sḵwx̱wú7mesh (Squamish), and Səl̓ílwətaʔ/Selilwitulh (Tsleil-Waututh) Nations, on whose traditional, unceded and ancestral territory we have the privilege of working.

Funding

This work was supported by Genome BC and Genome Canada (project # B26PMH), and Michael Smith Health Research BC (award #18932). The funders were not involved in the development of the study design, data analysis or interpretation, or writing of this manuscript.

Data availability

Data extracted from included RCTs are available from the corresponding author upon reasonable request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2023.115102.

References

Bansback, N., Tsuchiya, A., Brazier, J., et al., 2012. Canadian valuation of EQ-5D health states: preliminary value set and considerations for future valuation studies. PLoS One 7 (2), e31115. https://doi.org/10.1371/journal.pone.0031115.

Bousman, C.A., Arandjelovic, K., Mancuso, S.G., et al., 2019. Pharmacogenetic tests and depressive symptom remission: a meta-analysis of randomized controlled trials. Pharmacogenomics 20 (1), 37–47. https://doi.org/10.2217/pgs-2018-0142.

Bousman, C.A., Forbes, M., Jayaram, M., et al., 2017. Antidepressant prescribing in the precision medicine era: a prescriber's primer on pharmacogenetic tools. BMC Psychiatry 17 (1), 60. https://doi.org/10.1186/s12888-017-1230-5.

Bousman, C.A., Hopwood, M., 2016. Commercial pharmacogenetic-based decision-support tools in psychiatry. Lancet Psychiatry 3 (6), 585–590. https://doi.org/10.1016/s2215-0366(16)00017-1.

Brown, L.C., Stanton, J.D., Bharthi, K., et al., 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. https://doi.org/10.1002/cpt.2748.

Bradley, P., Shiekh, M., Mehra, V., et al., 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. https://doi.org/10.1016/j.jpsychires.2017.09.024.

Brown, L., Vranjkovic, O., Li, J., et al., 2020. The clinical utility of combinatorial pharmacogenomic testing for patients with depression: a meta-analysis. Pharmacogenomics 21 (8), 559–569. https://doi.org/10.2217/pgs-2019-0157.

Clinical Pharmacogenetics Implementation Consortium, 2022. Genes-Drugs. Updated 1 June 2022. https://cpicpgx.org/genes-drugs/ (accessed 11 January 2023).

Dunlop, B.W., Parikh, S.V., Rothschild, A.J., et al., 2019. Comparing sensitivity to change using the 6-item versus the 17-item Hamilton depression rating scale in the GUIDED randomized controlled trial. BMC Psychiatry 19 (1), 420. https://doi.org/10.1186/s12888-019-2410-2.

Evidence Prime Inc, 2020. GRADEpro GDT: GRADEpro Guideline Development Tool. McMaster University.

Forester, B.P., Parikh, S.V., Weisenbach, S., et al., 2020. Combinatorial pharmacogenomic testing improves outcomes for older adults with depression. Am. J. Geriatr. Psychiatry 28 (9), 933–945. https://doi.org/10.1016/j.jagp.2020.05.005.

Greden, J.F., Parikh, S.V., Rothschild, A.J., et al., 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. https://doi.org/10.1016/j.jpsychires.2019.01.003.

Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiat. 23 (1), 56–62. https://doi.org/10.1136/jnnp.23.1.56.

Han, C., Wang, S.M., Bahk, W.M., et al., 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. https://doi.org/10.9758/cpn.2018.16.1.469.

Han, C., Wang, S.M., Bahk, W.M., et al., 2020. Corrigendum: a Pharmacogenomic-based antidepressant treatment for patients with major depressive disorder: results from an 8-week, randomized, single-blinded clinical trial. Clin. Psychopharmacol. Neurosci. 18 (4), 641. https://doi.org/10.9758/cpn.2020.18.4.641.

Hicks, J.K., Bishop, J.R., Sangkuhl, K., et al., 2015. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for CYP2D6 and CYP2C19 genotypes and dosing of selective serotonin reuptake inhibitors. Clin. Pharmacol. Ther. 98 (2), 127–134. https://doi.org/10.1002/cpt.147.

Hicks, J.K., Sangkuhl, K., Swen, J.J., et al., 2017. Clinical Pharmacogenetics Implementation Consortium guideline (CPIC) for CYP2D6 and CYP2C19 genotypes and dosing of tricyclic antidepressants: 2016 update. Clin. Pharmacol. Ther. 102 (1), 37–44. https://doi.org/10.1002/cpt.597.

Ho, S.C., Jacob, S.A., Tangiisuran, B., 2017. Barriers and facilitators of adherence to antidepressants among outpatients with major depressive disorder: a qualitative study. PLoS One 12 (6), e0179290. https://doi.org/10.1371/journal.pone.0179290.

IsHak, W.W., Mirocha, J., James, D., et al., 2015. Quality of life in major depressive disorder before/after multiple steps of treatment and one-year follow-up. Acta Psychiatr. Scand. 131 (1), 51–60. https://doi.org/10.1111/acps.12301.

Knoll, A.D., MacLennan, R.N., 2017. Prevalence and correlates of depression in Canada: findings from the Canadian Community Health Survey. Can. Psychol./Psychologie Canadienne 58 (2), 116–123. https://doi.org/10.1037/cap0000103.

Kolovos, S., Bosmans, J.E., van Dongen, J.M., et al., 2017. Utility scores for different health states related to depression: individual participant data analysis. Qual. Life Res. 26 (7), 1649–1658. https://doi.org/10.1007/s11136-017-1536-2.

McGuinness, L.A., Higgins, J.P.T., 2021. Risk-of-bias VISualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. Res. Synth. Methods 12 (1), 55–61. https://doi.org/10.1002/jrsm.1411.

Menchon, J.M., Espadaler, J., Tuson, M., et al., 2019. Patient characteristics driving clinical utility in psychiatric pharmacogenetics: a reanalysis from the AB-GEN multicentric trial. J. Neural. Transm. (Vienna) 126 (1), 95–99. https://doi.org/10.1007/s00702-018-1879-z.

Ontario Health, 2021. Multi-gene pharmacogenomic testing that includes decision-support tools to guide medication selection for major depression: a health technology assessment. Ont. Health Technol. Assess. Ser. 21 (13), 1–214.

Oslin, D.W., Lynch, K.G., Shih, M.C., et al., 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. https://doi.org/10.1001/jama.2022.9805.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., et al., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372 (71). https://doi.org/10.1136/bmj.n71.

Perez, V., Salavert, A., Espadaler, J., et al., 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. https://doi.org/10.1186/s12888-017-1412-1.

Perlis, R.H., Dowd, D., Fava, M., et al., 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. https://doi.org/10.1002/da.23029.

Rosenblat, J.D., Lee, Y., McIntyre, R.S., 2017. Does pharmacogenomic testing improve clinical outcomes for major depressive disorder? A systematic review of clinical trials and cost-effectiveness studies. J. Clin. Psychiatry 78 (6), 720–729. https://doi.org/10.4088/JCP.15r10583.

Rosenblat, J.D., Lee, Y., McIntyre, R.S., 2018. The effect of pharmacogenomic testing on response and remission rates in the acute treatment of major depressive disorder: a meta-analysis. J. Affect. Disord. 241, 484–491. https://doi.org/10.1016/j.jad.2018.08.056.

Ruano, G., Robinson, S., Holford, T., et al., 2020. Results of the CYP-GUIDES randomized controlled trial: total cohort and primary endpoints. Contemp. Clin. Trials 89, 105910. https://doi.org/10.1016/j.cct.2019.105910.

Ruano, G., Tortora, J., Robinson, S., et al., 2021. Subanalysis of the CYP-GUIDES trial: CYP2D6 functional stratification and operational timeline selection. Psychiatry Res. 297, 113571 https://doi.org/10.1016/j.psychres.2020.113571.

Ruhe, H.G., Huyser, J., Swinkels, J.A., et al., 2006. Switching antidepressants after a first selective serotonin reuptake inhibitor in major depressive disorder: a systematic review. J. Clin. Psychiatry 67 (12), 1836–1855. https://doi.org/10.4088/jcp.v67n1203.

Rush, A.J., Trivedi, M.H., Wisniewski, S.R., et al., 2006a. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am. J. Psychiatry 163 (11), 1905–1917. https://doi.org/10.1176/appi.ajp.163.11.1905.

Rush, A.J., Trivedi, M.H., Wisniewski, S.R., et al., 2006b. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. N. Engl. J. Med. 354 (12), 1231–1242. https://doi.org/10.1056/NEJMoa052963.

Scherf-Clavel, M., Weber, H., Wurst, C., et al., 2022. Effects of pharmacokinetic gene variation on therapeutic drug levels and antidepressant treatment response. Pharmacopsychiatry. https://doi.org/10.1055/a-1872-0613.

Shan, X., Zhao, W., Qiu, Y., et al., 2019. Preliminary clinical investigation of combinatorial pharmacogenomic testing for the optimized treatment of depression: a randomized single-blind study. Front Neurosci. 13, 960. https://doi.org/10.3389/fnins.2019.00960.

Singh, A.B., 2015. Improved antidepressant remission in major depression *via* a pharmacokinetic pathway polygene pharmacogenetic report. Clin. Psychopharmacol. Neurosci. 13 (2), 150–156. https://doi.org/10.9758/cpn.2015.13.2.150.

Sterne, J.A.C., Savovic, J., Page, M.J., et al., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 366, l4898. https://doi.org/10.1136/bmj.l4898.

Swen, J.J., Nijenhuis, M., de Boer, A., et al., 2011. Pharmacogenetics: from bench to byte-an update of guidelines. Clin. Pharmacol. Ther. 89 (5), 662–673. https://doi.org/10.1038/clpt.2011.34.

Swen, J.J., Wilting, I., de Goede, A.L., et al., 2008. Pharmacogenetics: from bench to byte. Clin. Pharmacol. Ther. 83 (5), 781–787. https://doi.org/10.1038/sj.clpt.6100507.

Tansey, K.E., Guipponi, M., Hu, X., et al., 2013. Contribution of common genetic variants to antidepressant response. Biol. Psychiatry 73 (7), 679–682. https://doi.org/10.1016/j.biopsych.2012.10.030.

Thase, M.E., Parikh, S.V., Rothschild, A.J., et al., 2019. Impact of pharmacogenomics on clinical outcomes for patients taking medications with gene-drug interactions in a randomized controlled trial. J. Clin. Psychiatry 80 (6). https://doi.org/10.4088/JCP.19m12910.

Tiwari, A.K., Zai, C.C., Altar, C.A., et al., 2022. Clinical utility of combinatorial pharmacogenomic testing in depression: a Canadian patient- and rater-blinded, randomized, controlled trial. Transl. Psychiatry 12 (1), 101. https://doi.org/10.1038/s41398-022-01847-8.

van der Schans, J., Hak, E., Postma, M., et al., 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. https://doi.org/10.1097/JCP.0000000000001129.

Vande Voort, J.L., Orth, S.S., Shekunov, J., et al., 2021. A randomized controlled trial of combinatorial pharmacogenetics testing in adolescent depression. J. Am. Acad. Child Adolesc. Psychiatry 61 (1), 46–55. https://doi.org/10.1016/j.jaac.2021.03.011.

Winner, J.G., Carhart, J.M., Altar, C.A., et al., 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.

World Health Organization, 2017. Depression and Other Common Mental Health Conditions. WHO Document Production Services, Geneva, Switzerland, pp. 1–24.