Pharmacogenomic Clinical Support Tools for the Treatment of Depression

Matthew L. Baum, M.D., Ph.D., Alik S. Widge, M.D., Ph.D., Linda L. Carpenter, M.D.,
William M. McDonald, M.D., Bruce M. Cohen, M.D., Ph.D., Charles B. Nemeroff, M.D., Ph.D.,
On behalf of the American Psychiatric Association (APA) Workgroup on Biomarkers and Novel Treatments

Objective: In this review, the authors update the 2018 position statement of the American Psychiatric Association Council of Research Workgroup on Biomarkers and Novel Treatments on pharmacogenomic (PGx) tools for treatment selection in depression.

Methods: The literature was reviewed for new clinical trials and meta-analyses, published from 2017 to 2022, of studies using PGx tools for treatment selection in depression. The blinding and control conditions, as well as primary and secondary outcomes and post hoc analyses, were summarized.

Results: Eleven new clinical trials and five meta-analyses were identified; all studies had primary outcome measures related to speed or efficacy of treatment response. Three trials (27%) demonstrated efficacy on the primary outcome measure with statistical significance; the three studies used different PGx tools; one study was open-label and the other two were small single-blind trials. Five trials (45%) did not detect efficacy with statistical significance on either primary or secondary outcome measures. Only one trial (9%) used adverse events as a primary outcome measure. All studies had significant limitations; for example, none adopted a fully blinded study design, only two studies attempted to blind the treating clinician, and none incorporated measures to estimate the effectiveness of the blinds or the influence of lack of blinding on the study results.

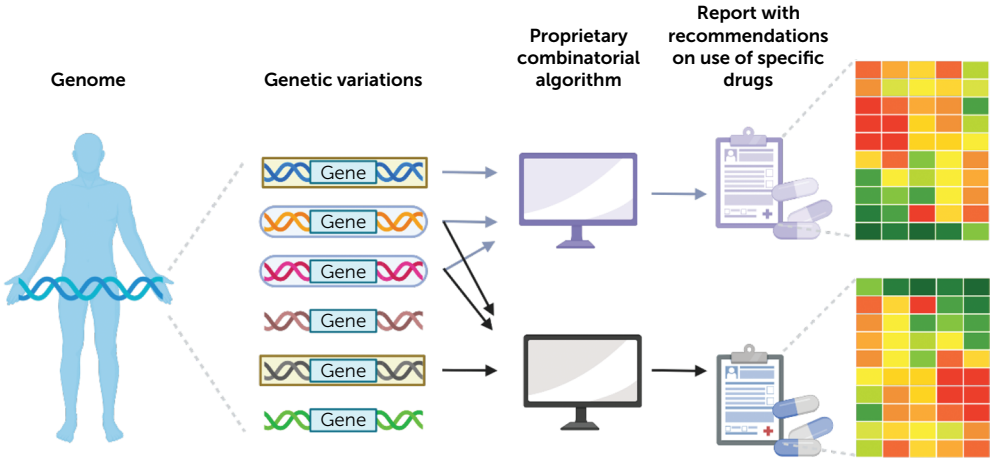
Conclusions: The addition of these new data do not alter the recommendations of the 2018 report, or the advice of the U.S. Food and Drug Administration, that the evidence does not support the use of currently available combinatorial PGx tools for treatment selection in major depressive disorder. Priority efforts for future studies and the development and testing of effective tools include fully blinded study designs, inclusion of promising genetic variants not currently included in any commercially available tests, and investigation of other uses of pharmacogenomics, such as estimating the likelihood of rare adverse drug effects, rather than increasing the speed or magnitude of drug response.

Am J Psychiatry 2024; 181:591–607; doi: 10.1176/appi.ajp.20230657

There is intense interest in whether pharmacogenomic (PGx) clinical support tools might be useful to improve treatment outcomes in patients with depression. In 2018, the American Psychiatric Association Council on Research (APA-COR) Workgroup on Biomarkers and Novel Treatments reviewed the available published data and concluded that there was insufficient evidence to support the widespread use of PGx tools in clinical practice (1). The position statement also suggested useful future directions for PGx research. The U.S. Food and Drug Administration (FDA) subsequently issued a warning letter (2) as part of its "ongoing efforts to protect the public from the significant risk these tests pose." The FDA news release on the 2019 letter indicated that the FDA was "particularly concerned about pharmacogenetic tests that claim to predict patients' responses to specific medications where such claims have not been established and are not described in the drug labeling" and that the FDA continues "to warn patients and health care professionals that they should not rely on these tests for treatment decisions." That same year, the International Society of Psychiatric Genetics published an updated policy statement that also did not encourage widespread use of PGx, emphasizing that testing for CYP2C19 and CYP2D6 was most appropriate for patients who had "experienced inadequate response or adverse reactions to a previous antidepressant" (3). Despite expert opinions, warnings, and policy statements regarding their limitations for predicting antidepressant treatment response, the popularity of PGx testing products has grown, with at least 35 U.S. commercial entities providing them by 2020 (4). Since the APA-COR workgroup's 2018 review paper, nearly a dozen additional randomized PGx controlled trials have been completed. Here,

See related features: **CME course** *(online and p. 607) and* **Video** *by Dr. Pine (online)*

FIGURE 1. Understanding pharmacogenomic (PGx) tools in psychiatry[a]



[a] *Genome*: Each individual person has a genome containing a large number of genetic variants that impact the kinetics and dynamics of drugs that may be prescribed to treat them. *Genetic variations*: PGx tools perform assays on a subset of the relevant genetic variants to determine the individual's genotype for those specific genes; all PGx tools assay a certain subset of genetic variants (genes shown outlined by ellipses), while other subsets of genetic variants are tested by some but not all pharmacogenetic tools (genes outlined by rectangles). A third subset of important genetic variants are not currently assayed by any commercially available PGx tools (genes with no outlines). *Proprietary combinatorial algorithm*: Each company's tool uses its proprietary algorithm to weight and combine the effect of the variants it assays to produce their results report. *Report with recommendations on use of specific drugs*: The main output is a table of specific drugs with recommendations on how to use them, such as "use as directed," "moderate gene-drug interaction," and "severe gene-drug interaction," as represented by the colors in the table here. Notably, recommendations from different PGx tools have been shown to differ from each other substantially. Figure created using BioRender (biorender.com).

the workgroup reviews this additional body of research to provide updated recommendations.

The literature was reviewed for new clinical trials and meta-analyses, published from 2017 to 2022, of studies using PGx tools for treatment selection in depression. Eleven new clinical trials and five meta-analyses were identified. The blinding and control conditions, as well as primary and secondary outcomes and post hoc analyses, were summarized.

WHAT ARE PHARMACOGENOMIC CLINICAL SUPPORT TOOLS?

Currently available PGx clinical support tools use information about genetic variants to advise on selection of antidepressants (see Figure 1).

There are two potential mechanisms by which PGx tests can contribute to predicting antidepressant drug effectiveness: genetic variants that affect pharmacokinetics and genetic variants that affect pharmacodynamics. Pharmacokinetics refers to how much of a drug gets into circulation; how it distributes among organs and tissues, thus determining the quantity that enters the brain; and how fast the drug is inactivated and eliminated from the body, thus determining how long it circulates in the body. Genetic variants that affect the metabolism, tissue distribution, inactivation, and elimination of antidepressants affect pharmacokinetics. The majority of known associations between genotype and blood concentrations of antidepressant drug are related to cytochrome P450 hepatic enzymes. Pharmacodynamics refers to the biological targets through which a drug produces its effects. For example, genetic variants that affect the affinity of a drug for its target receptor(s) may influence the effect of a drug.

Treatment outcomes in depression are dependent on both pharmacokinetics and pharmacodynamics, and most PGx clinical support tools provide information on a panel of individual genetic variants. Most of the pharmacokinetic genetic variants involve testing for the cytochrome P450 enzymes (e.g., CYP2D6, CYP3A4). The current commercial PGx clinical support tools differ from each other in terms of the number and identity of genes included in the tool, and often, which genetic variants are tested. The rationale for why certain variants are included is often not disclosed.

The pharmacokinetics and pharmacodynamics of antidepressant medications are affected by variants at multiple genes. The final recommendations by PGx clinical support tools come from a combinatorial algorithm that weighs the impact of each genetic variant. The algorithms used by each company, like the choice of genes included for testing, are usually proprietary, making it impossible for prescribers or consumers to determine and evaluate the validity of decisions made by the algorithms and their translation to recommendations about a patient's pharmacotherapy. Although there is some variability in the content of reports generated by different PGx clinical support tools, most reports consist of both a description of the genes assayed and specific recommendations for treatment with medications commonly prescribed for patients with depression—for example, "use as directed," "use with caution," or "use with increased caution and with more frequent monitoring." Often the report is color coded to represent the relative salience level of each recommendation for the specific patient tested. While these reports are mainly structured to provide information that can inform drug dosing (e.g., whether a lower dosage might be needed in a poor metabolizer to avoid adverse events), they are often regarded by patients and clinicians as guides for predicting the probability of clinical response.

Given that the various PGx tools differ in the specific genes, variants, and types of analyses used to process patient-specific genetic information and provide clinical recommendations, it is not surprising that recent work (5) has highlighted the poor concordance of treatment recommendations provided across PGx clinical support tools from

different companies. In other words, the tools are far from interchangeable with one another, and evidence for or against the usefulness of one of them does not directly translate to the others.

WHY HAVE PGx CLINICAL SUPPORT TOOLS BEEN INVESTIGATED FOR USE IN GUIDING TREATMENT OF DEPRESSION?

One rationale put forward for the potential usefulness of PGx clinical support tools is that antidepressant medication prescribing is a process of trial and error, often requiring serial trials of multiple agents (either as monotherapy or combination therapies), and each trial may require up to 8 weeks. For example, in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, slightly more than one-quarter of patients achieved remission with the first antidepressant (citalopram), and one-quarter of those who stayed in the study for a second antidepressant trial achieved remission, and one-eighth of those who stayed in the study for a third antidepressant trial achieved remission (6). Because genetic factors have been suggested to account for a large part of the variation in antidepressant response (e.g., 7), it was hypothesized that guiding drug selection using an evaluation of variants in the genome would lead to less time lost to the trial-and-error process and better response, remission, and tolerability. It is worth noting, however, that commercially available PGx tools include only a subset of the known potentially relevant genomic variants. Some genetic variants with potential relevance, such as the norepinephrine transporter and the corticotropin-releasing hormone binding protein, are not included in any of the current tools (8–11). Furthermore, the strength of associations between gene variants and clinical response has not been established, nor is it known how many and which genes are the most relevant in determining depression symptom response.

WHY IS AN UPDATED RECOMMENDATION ON THE APPROPRIATE USE OF PGx CLINICAL SUPPORT TOOLS NEEDED?

In the 2018 APA-COR position statement (1), we highlighted that 1) the scientific rationale behind the collection of genes chosen for testing in the commercial tools was inadequate, 2) no randomized controlled trial of any PGx clinical support tool had demonstrated efficacy on its primary clinical endpoint measure, and 3) all studies contained serious methodological weaknesses, such as lack of proper or any blinding, appearance of conflicts of interest, and small sample sizes. Based on this combination of concerns, we concluded that there was insufficient evidence to support the widespread use of combinational PGx clinical support tools in the treatment of depression.

Others have subsequently published critiques of PGx depression studies (see Box 1). Notably, the FDA released a position statement strongly advising against the use of these tools in the treatment of depression (12, 13). In 2020, the FDA released an updated communication (14) reiterating its prior warning about the clinical use of PGx tools. It stated that "Unfortunately, in the time since our safety communication was issued, some manufacturers of pharmacogenetic tests with claims not adequately supported by sound science have continued marketing their tests, including some for medications to treat seizures, mental illness, and pain, including opioids. The FDA remains concerned with the safe use of these medications based on pharmacogenetic test reports that are not supported by sound science." The statement explained the potential consequences of using these tests: "Decisions based on inaccurate information can result in patient harm because patients may not receive the most appropriate medication, may receive a medication that could be harmful, or may receive a prescription for an inappropriate dose. All of these scenarios can create unnecessary delays and prevent patients from receiving the most timely and appropriate treatment. Patients and clinicians deserve better."

The FDA did specifically note that a very limited set of gene-drug interactions do have sufficient evidence to be clinically helpful, and it released new web-based resources listing specific drugs for which the evidence does support genotyping as a possible aid in patient care. The FDA has incorporated its guidance on individual cytochrome P450 gene variants into the labels of some antidepressants so that clinicians are aware that some patients may do better on adjusted dosages, and it has published a table with its guidelines (15). The Clinical Pharmacogenetics Implementation Consortium has proposed a similar set of guidelines (16). The FDA table separates recommendations into three categories. The first category reports gene-drug interactions that the FDA finds "support therapeutic management recommendations"; these are very limited and include only four medications typically used for depression: citalopram (for higher risk of QTc prolongation in CYP2D6 poor metabolizers), venlafaxine (considering dosage reductions in CYP2D6 poor metabolizers), and the atypical antipsychotics aripiprazole and brexpiprazole (considering dosage reductions in CYP2D6 poor metabolizers). The second category lists "potential impact on safety or response" and contains no common antidepressants. The third category lists gene-drug interactions that "affect pharmacokinetics only" and contains many common antidepressant agents, and the FDA specifically emphasizes that "the impact of these genetic variants or genetic variant inferred phenotypes on the safety or response of the corresponding drug has not been established" (15). All PGx tools currently make recommendations beyond just citalopram, venlafaxine, brexpiprazole, and aripiprazole and imply that their combinatorial algorithms incorporate information beyond CYP2D6 poor-metabolizer status.

Importantly, work on pharmacogenomics in other areas of medicine has taught us that even when individual genetic variants can have a large effect on drug levels, testing and changing treatments based on that information does not

BOX 1. Conclusions and concerns arising from expert reviews of pharmacogenomic (PGx) tools

- Studies have marked methodological limitations, including inadequate sample sizes, lack of blinding, industry influence, lack of gold-standard control conditions like clinical best practice guidelines, and failure to convincingly show effectiveness (1, 20, 69, 70).
- The FDA warned that "the relationship between DNA variants and the effectiveness of antidepressant medications has never been established. Moreover, . . . changes to patients' medication based on genetic test results that claim to provide information on the personalized dosage or treatment regimens for some antidepressant medications . . . could potentially lead to patient harm" (see reference 12). The FDA issued the first warning in 2018 (13) and updated it in 2020 (14).
- The role played by common genetic variants in expression and activity of drug metabolic pathway enzymes is likely minor compared to other factors (e.g., age, sex, extent of xenobiotic induction or inhibition, hormonal state) (66, 71, 72).
- PGx tools are more likely to recommend newer drugs (that have fewer known gene-drug interactions) or drugs that are not metabolized by the cytochrome P450 pathway (information that a treating clinician would know or could learn without a test) (73).

always lead to a clinical benefit. As an example (to which we will return in the Discussion section), even though it is well established that CYP2C19 is necessary to convert the antiplatelet agent clopidogrel from a prodrug to its active form, testing for CYP2C19 loss of function and switching to an alternative agent did not lead to a clinical benefit in large randomized trials (17, 18).

Below, we review the new clinical trials and meta-analyses on PGx tools in the treatment of depression. We review trials for each PGx tool, then review the meta-analyses. We include key features about the trials in tabular format in Tables S1–S4 in the online supplement.

THE STATE OF EVIDENCE OF PGx CLINICAL SUPPORT TOOLS IN THE TREATMENT OF DEPRESSION

GeneSight

The GeneSight pharmacogenomic combinatorial tool has been the most extensively studied and has had trials in both adults and adolescents. As most of the trials were industry funded (see Tables S1 and S2 in the online supplement for details on disclosures in the studies), we paid special attention to shared design choices that could impact results, and found that none of the studies appropriately blinded the treating clinicians.

GUIDED Trial and Subanalyses

Greden et al., 2019 (the GUIDED trial). The largest partially blinded study to date is the GUIDED trial, an 8-week randomized controlled trial (RCT) with a 16-week open-label follow-up of 1,167 patients with a diagnosis of major depressive disorder (MDD) who had at least one prior antidepressant trial (19). The study was conducted across 60 clinical sites in primary care and psychiatric practices. A majority of the sample were female (~70%) and Caucasian (~80%). The study compared treatment selection guided by GeneSight with treatment as usual (TAU). Clinicians had access to the PGx report and knew whether a patient was in the PGx-guided arm or in the TAU comparison group, although patients and raters performing clinical outcome assessments were both blinded. Clinicians were not required to prescribe according to the PGx report. The study was originally designed with unblinding after week 12, but because treating clinicians received PGx reports for patients in the TAU arm after week 8 and the study team was concerned that treatment personnel may have disclosed this to the TAU arm, only outcomes up to week 8 were considered "blinded" for the purposes of analysis. Analysis was performed with an intent-to-treat sample (N=1,541). The effectiveness of the blind was not assessed.

The primary efficacy endpoint measure was symptom improvement on the 17-item Hamilton Depression Rating Scale (HAM-D). The study did not find a significant difference between the PGx-guided group and the TAU group on this measure at week 8 (27.2% vs. 24.4% reduction in HAM-D score, $p=0.107$). Secondary outcomes were categorical response and remission; response was defined as a decrease ≥50% in HAM-D score, and remission was defined as a score ≤7 on the HAM-D. Secondary outcomes revealed statistically significantly higher rates of response and remission in PGx-guided treatment (response: 26% vs. 19.9%, $p=0.013$; remission: 15.3% vs. 10.1%, $p=0.007$). The study did not discuss why both the PGx-guided remission rates and the TAU remission rates were lower than in TAU depression trials like STAR*D, where remission rates were ~30% for the first and second treatment steps, and ~13% for the third and fourth treatment steps (6).

In both the PGx-guided and TAU arms, the study tracked whether participants' prescribed medications were considered "congruent" with recommendations for them in their PGx reports (i.e., "use as directed" [no gene-drug interactions] or "use with caution" [moderate gene-drug interactions]) versus "incongruent" (i.e., the patient's prescribed agents were among those labeled "use with increased caution and with more frequent monitoring" [significant gene-drug interactions]). At baseline, 81.7% of participants were on medications that were considered congruent with their report's recommended use profile. Consistent with expectations that clinicians would often follow the advice they received

in the reports, use of "congruent" medications increased from 79.4% to 91.2% in the PGx-guided arm but stayed at approximately 80% in the TAU arm. In a post hoc subgroup analysis of rates of side effects in the group of patients on PGx-"congruent" medications or PGx-"incongruent" medications, they found no significant difference (p=0.855) in side effects.

A post hoc subgroup analysis focused on participants taking "incongruent" medications at baseline (N=213) and those who switched to "congruent" medications during the trial (N=136), regardless of treatment arm. The subgroup with "incongruent" medications at baseline that switched to "congruent" medications had a larger decrease in HAM-D score (33.5% vs. 21.1%, p=0.002) than the subgroup that stayed on "incongruent" medications (N=77). Additionally, at week 8, patients on "congruent" medications showed higher rates of response (28.5% vs. 16.7%, p=0.036) and remission (21.5% vs. 8.5%, p=0.007). Those who switched to "congruent" medications from "incongruent" medications also had lower mean numbers of side effects (6.5% vs. 16.5%, p=0.045). While these post hoc results provide some support for the potential utility of the test that would merit follow-up in a trial as a primary outcome, the published results did not test plausible alternative explanations for these results. For example, in past studies, some of the "incongruent" medication choices included uncommon agents with poorer baseline tolerability or efficacy that consequently were not concordant with standard treatment of depression (20). That is, a switch from an inappropriate to an appropriate antidepressant treatment would be expected to be beneficial in many cases, and the test result helped encourage a switch to standard treatment. In addition, as documented by STAR*D, a medication switch often leads to some improvement. The issue for study is whether PGx leads to a better choice than standard protocol guidance in such choices. The authors did not quantify the effect of changes to "congruent" medications that do not require a PGx test to know that they would be congruent (e.g., desvenlafaxine, which is not metabolized by cytochrome P450 enzymes or affected by other variants assayed in the test, was used at a frequency of 1.8% at baseline, and the rate increased to 8.4% in the PGx-guided group but stayed at 2.5% in the TAU group). The authors also did not conduct a power analysis to estimate how many patients would have to be tested to see an effect on their primary outcome.

Subanalysis in GUIDED trial: effect on those with drug-gene interactions at baseline (Thase et al., 2019). This follow-up report described a secondary analysis of the GUIDED trial outcome data, examining the subpopulation of patients (N=912) who had any predicted drug-gene interactions at baseline, whether moderate or significant. The original GUIDED trial analysis did not quantify numbers or outcomes for patients who started on medications in the "use with caution" category, as this category was pooled together with "use as directed" and classified as "congruent" in the primary analysis. The 2019 analysis by Thase et al. (21) found that compared with patients in the TAU group, patients with gene-drug interactions at baseline who ended up in the PGx-guided arm had greater symptom improvement on the HAM-D at 8 weeks (27.1% vs. 22.1% reduction, p=0.029), a higher response rate (although it fell short of significance: 27% vs. 19%, p=0.08), and significantly higher remission rates (18.2% vs. 10.7%, p=0.003). In a second analysis that looked at only the subset of patients with baseline drug-gene interactions who switched medications during the trial (N=460), group differences were also significant. Among patients in the PGx-guided and TAU arms, respectively, the rates of symptom improvement were 30% and 22.3% (p=0.011), response rates were 29.8% and 19.4% (p=0.011), and remission rates were 20.3% and 11.1% (p=0.008). The authors did not quantify whether outcomes differed between patients who switched to medications without gene-drug interactions and those who switched to medications with drug-gene interactions.

Because the treating clinicians were not blinded to the treatment arm, there is the possibility that the treating clinicians treated the groups differently in a way that could affect outcome separately from any value of the PGx test results. Subanalyses that could have provided clues to the presence and magnitude of this bias were not conducted. For example, one might predict that if gene-guided medication switches are the mechanism of benefit in the PGx arm, then a) patients in the PGx arm who switched to medications without gene-drug interactions should have more benefit than those in the PGx arm who switched to a medication that still had gene-drug interactions, and b) those in the PGx arm who did not switch medications should not have more benefit than those in the TAU arm who did not switch medications. These comparisons were not done, even though they would have been similarly or better powered than the incongruent-versus-congruent comparison subanalysis in the original trial report (in which the Ns were 136 and 77). The authors quantified the percentage of switches in the PGx arm that resulted in patients being on a medication with no predicted gene-drug interaction (66.4%), and thus the Ns for this comparison would be ~152 for no gene-drug interaction and 83 for gene-drug interaction. For the second comparison, of those with medications with gene-drug interactions at baseline in the PGx and TAU groups who did not switch, the Ns were 122 and 205, respectively.

Forester et al., 2020. Forester et al. also published results of a secondary analysis of data from the GUIDED trial, using a subset of the sample age 65 and older (22). Analysis of an intent-to-treat sample, with no correction for performing a secondary analysis of data, found no statistically significant difference in symptom improvement at week 8 for the PGx-guided group versus the TAU group on the study's primary outcome measure (symptom improvement on the HAM-D) among these older participants; however, response and remission rates were both superior at week 8 for the PGx-guided group (response: 29.6% vs. 16.1%, p=0.032;

remission: 20.1% vs. 7.4%, p=0.014). There was no significant difference between groups in adverse drug events (10.2% vs. 7%, p=0.435). The authors also plotted the distribution of change from baseline to 8 weeks on the HAM-D for participants in the PGx-guided and TAU arms; visual inspection revealed a continuous distribution for the TAU arm and a possible bimodal distribution in the PGx-guided arm, suggesting a greater magnitude of symptom reduction among responders in the PGx-guided arm compared with the TAU arm (no statistical comparisons were performed).

PRIME Care Trial

Oslin et al., 2022. The largest randomized single-blind (rater only) trial of a PGx tool was the PRIME Care trial (23, 24). It was a 24-week pragmatic trial of 1,944 veterans with MDD and at least one prior treatment episode. Notably, this trial's sex ratio was the opposite of the ratio of most previous trials, with ~25% women (most other trials had samples in which a majority were women). To help estimate sample size, the results of a previous study (25) were used to estimate that approximately 20% of this population would have a next intended antidepressant with a clinically significant gene-drug interaction (that is, when clinicians were asked what medication they would think to switch the patient to next, without use of genetic information, the medications they chose would have a gene-drug interaction in about 20% of cases and no significant gene-drug interaction in about 80% of cases). The patients and treating clinicians knew whether they were in the PGx-guided or TAU arms, but assessments were conducted by telephone by a centralized call center of blinded raters. The study compared treatments guided by GeneSight versus TAU, and outcomes were measured at 4, 8, 12, 18, and 24 weeks. The study had two primary endpoints: number of prescriptions with predicted gene-drug interactions and rate of remission, defined as a score ≤5 on the Patient Health Questionnaire–9. The study met the primary endpoint of number of prescriptions with predicted gene-drug interactions, with the PGx group having 59.3% with no gene-drug interaction, 30.0% with moderate gene-drug interaction, and 10.7% with substantial gene-drug interaction, compared with 25.7%, 54.6%, and 19.7%, respectively, in the TAU group; the differences were statistically significant in a proportional logistic model. These results suggest that the clinicians followed the advice received from the tool but do not assess clinical utility.

For the second primary endpoint, remission, the study used a more complex group-by-time interaction model, integrating data across the 24 weeks. There was a significant effect of group (odds ratio=1.28, 95% CI=1.05–1.57, p=0.02; absolute risk difference, 2.8%, 95% CI=0.6–5.1) but not of group by time (p=0.08). The authors note significant differences between groups at weeks 8 and 12 but not the other time points. It is unclear, therefore, how to interpret the nature or durability of this effect. Secondary outcomes of response and symptom reduction were also met, but the authors point out that there was no correction for multiple hypothesis testing and indicate that the methods they used should thus be considered exploratory (odds ratio=1.25, 95% CI=1.07–1.46, p=0.005; absolute risk difference, 4.0%, 95% CI=1.2–6.8, and mean difference in reduction between groups, 0.56, 95% CI=0.17–0.95, p=0.005). Although there was no significant intervention-by-covariate interaction, a higher proportion of patients in the PGx group received a prescription for an antidepressant medication in the first 30 days. The study did not examine differential rates of adverse drug events. Most importantly, the first primary outcome measure fundamentally represents a tautology: the PGx-guided group, by definition, would have fewer gene-drug interactions if the clinicians followed the advice received from the tool. No subanalyses were performed to investigate the value of a switch to a medication without predicted gene-drug interactions, or to investigate the presence of an effect due to the lack of blinding (e.g., a placebo-like effect if there was a hope that gene-guided decisions would be more likely to be helpful than TAU decisions).

GAPP-MDD

Tiwari et al., 2022. The GAPP-MDD trial was a 52-week, three-arm RCT of 371 patients with MDD who had inadequate response to at least one prior medication trial during the current episode. All 371 participants were included in the intent-to-treat analysis, and 276 were included in the per protocol analysis (in which patients with no or mild depression at time of assessment, as assessed with the 17-item HAM-D, were excluded). Approximately 60% of the sample were female. The study compared treatment selection guided with GeneSight versus TAU (26). There were three trial arms, which included two interventional arms (one guided with regular GeneSight alone, and another where the GeneSight report was "enhanced with six gene variants associated with antipsychotic-induced weight gain") and a TAU arm (in which clinicians did not receive a PGx report). The rater and patient were blinded, but the treating clinician was aware of assignment. Analysis of outcomes was prespecified to combine both interventional arms into one PGx-guided arm if no difference was observed between interventional arms, and this was indeed done. With the combined PGx-guided pool, the primary outcome was change in HAM-D score in the per protocol analysis at week 8, which was 27.7% in the PGx-guided group and 22.7% in the TAU group and did not attain statistical significance (p=0.274) (reductions in HAM-D scores were 5.2 and 5.1 points, respectively; p=0.901). Similar results were obtained in the intent-to-treat protocol. Similarly, the secondary outcomes of response and remission at 8 weeks (response: 30.3% vs. 22.7%, p=0.262; remission: 15.7% vs. 8.3%, p=0.131) and at 24 weeks were not significantly different between treatment groups.

The authors noted that the power of this trial to find the small effect sizes observed in the GUIDED trial would have only been 25%, so it was underpowered. They hypothesized this was due in part to ~80% of people taking "gene-congruent"

medications at baseline and the observation that PGx increased that proportion by only approximately 8%; they found no statistically significant difference in percentage of gene-congruent medications at week 8 between groups (p=0.07).

The authors also performed a meta-analysis combining this trial, the GUIDED trial (19), and a 2013 study by Winner et al. (27) and reported a decrease of an additional 3.33% in HAM-D score in the PGx-guided arm compared with the TAU arm at week 8 or 10 (p=0.039) and odds ratios of 1.44 for response (p=0.004) and 1.69 for remission (p=0.001). They did not include the 2012 and 2013 open-label GeneSight studies by Hall-Flavin et al. (28, 29).

GeneSight in Adolescent Mental Health

Vande Voort et al., 2022. This was an 8-week RCT with 6-month follow-up for 179 adolescents (ages 13–18) with moderate to severe MDD (30), of whom 155 completed the full 8-week trial. The study population was remarkable in the landscape of PGx studies because it was an adolescent population. It was similar to the adult studies in that the majority of participants were female (~78%). The authors compared treatment selection guided by GeneSight versus TAU. Raters and patients/families were blinded, but treating clinicians were not. The primary outcome was difference on the Children's Depression Rating Scale (CDRS), and it did not reach statistical significance; at 8 weeks, there was a 30.7% decrease in the PGx-guided arm, compared with a 29.1% decrease in the TAU arm (p=0.889). There was no statistically significant difference in secondary outcomes of response or remission between the two treatment groups. This trial was designed to have power of 90% to detect an expert-defined minimal clinically important difference of 4 in CDRS score, and ended up being powered at 80%, although the authors suggest it should still be interpreted as a negative rather than an underpowered trial. Scores on the Frequency, Intensity, and Burden of Side Effects Rating Scale (FIBSER) were not significantly different between groups (p=0.28 at 8 weeks). There was no difference in patient/family satisfaction. In the PGx-guided group, 91.7% of patients had clinicians who chose medications based on the PGx report. The PGx-guided physicians were statistically less likely than those in the TAU arm to prescribe selective serotonin reuptake inhibitors (SSRIs) and more likely to prescribe serotonin-norepinephrine reuptake inhibitors and atypical antidepressants. Given that the only FDA-approved medications for depression in adolescence are SSRIs (fluoxetine and escitalopram), this suggests that PGx can change physician prescribing toward less evidence-based practice. This choice illustrates the FDA's stated concern that use of genetic tools at this time can lead to potentially inappropriate clinical decisions. Also, the authors note that desvenlafaxine, which is not metabolized by or expected to be impacted by variants of the tested genes, is consistently on the reports as "use as directed" even though the report by definition provides no useful information on that drug.

Neuropharmagen

Neuropharmagen is a PGx tool used most frequently outside of North America. It has been the subject of the studies summarized below, not reviewed in our last evaluation. The output of the tool is a report that includes color coding to designate medications' gene-drug interactions; green indicates an "increased likelihood of positive response and/or lower risk of adverse drug reactions"; red indicates an "increased risk of adverse drug reactions"; yellow indicates a "need for drug dose monitoring and/or less likelihood of positive response"; and white indicates "no genetic variants relevant to the treatment have been found; use as directed" (31).

AB-GEN trial (Pérez et al., 2017). This was a 12-week RCT of 316 patients with MDD, ~85% of whom had at least one medication trial with insufficient improvement (32). The majority of the patients (64%) were female. Treating clinicians were not blinded to PGx reports, but raters and patients were. The study compared PGx-guided treatment selection versus TAU. The primary endpoint was proportion of patients achieving a sustained response, as indicated by a score ≤2 on the Patient Global Impression of Improvement scale (PGI-I) for at least two measurement periods (e.g., at 8 and 12 weeks). Statistical significance at the primary endpoint was not met (38.5% vs. 34.4%, p=0.474; odds ratio=1.19, 95% CI=0.74–1.92). A range of subanalyses were performed, and a nominally higher rate of PGI-I scores ≤2 was observed at 12 weeks in the PGx group (47.8% vs. 36.1%, p=0.048; odds ratio=1.62, 95% CI=1.00–2.61), but the significance level did not meet the threshold calculated for correction for multiple comparisons (a p value <0.027). Side effects, however, as assessed by FIBSER scores, did surpass the significance threshold calculated for correction for multiple comparisons. The study was likely underenrolled, as the investigators estimated their desired power at N=520 with 25% dropout, for a target N of 390. A secondary analysis (33) found significant effects for their outcome measures if only the patients with moderate or severe depression were analyzed, or if only patients under age 60 were analyzed.

Han et al., 2018. This was a prospective RCT in Korea of 100 patients with treatment-refractory depression (31). The study population is of note because it is among the few outside of Europe or North America, and the mean duration of illness was 6 years, with at least two prior ineffective medication trials. The population was majority female (~75%), similar to most other studies reviewed here. The trial compared treatment selection using Neuropharmagen versus TAU. Only the patient was blinded to treatment. The primary endpoint was decrease in score on the 17-item HAM-D, which was reported to be statistically significant in favor of PGx (−16.1 [SD=6.8] vs. −12.1 [SD=8.2], p=0.010 via analysis of covariance). Among secondary endpoints, response was also found to be statistically significant in favor of PGx (71.7% vs. 43.6%, p=0.014), but the difference in

remission rates was not statistically significant (45.5% vs. 25.6%, p=0.071). The authors acknowledge the potential for observation bias given that clinicians/raters were not blinded, and they do not comment on their very high rates of response and remission given this treatment-refractory population.

OTHER PGx TOOLS

Since our last review, there have been single trials of five other tools, reviewed below: Mental Health DNA Insight; NeuroIDgenetix; Pillcheck; Genecept, version 2.0; and a custom algorithm.

Mental Health DNA Insight

McCarthy et al., 2021. This was a double-blind RCT of 182 veterans with transdiagnostic treatment-resistant depression (34) (defined as having had at least one adequate prior medication trial for treatment of a depressive episode in the context of MDD, bipolar depression, or posttraumatic stress disorder [PTSD]). The study used Mental Health DNA Insight, from Pathway Genomics, which "placed 53 medications into one of four use categories: preferential use, use as directed, may have significant limitations, and may cause serious adverse events." This was the only recent PGx trial in which both the treating clinicians and the patients were blinded. For patients in the TAU group, the treating clinician was provided a sham PGx report that listed all the medications as "use as directed," which is not the format of the typical PGx report. In addition, raters in this study were also the treating clinicians, so the study did not have a fully blinded design. The Clinical Global Impressions severity scale (CGI-S) was the primary endpoint. Remission was defined as a CGI-S score of 1 or 2. Statistical significance was not met for the primary endpoint (p=0.8). There were also no significant differences between groups in adverse events. Subanalysis found statistically significant improvement on the CGI-S for PGx over TAU among patients with comorbid PTSD (p<0.001, which survived Bonferroni correction). Of note, the investigators did not ask clinicians or patients to guess which arm of the trial a given patient was in, and thus there is no measure of how successful the blinding was. Because a PGx report in which all drugs are listed as "use as directed" would only be found if a person had the drug-neutral genotype at every locus assayed in the genomic tool, and thus perhaps rare, it would be important to know if the sham report could be easily detected as a sham report.

NeuroIDgenetix

This PGx tool, in addition to reporting results of pharmacokinetic and pharmacodynamic variants, also screens for drug-drug interactions and incorporates lifestyle factors, including substance and herbal supplement use. Medications are classified as "use as directed" or "use with caution and/or increased monitoring," with a brief description of the reasons—for example, "lack of efficacy or toxicity."

Bradley et al., 2018. This was a 12-week RCT of 579 patients with depression and anxiety (35). Patients were predominantly female (73%), and the proportions of those with depression, anxiety, and depression plus anxiety were roughly balanced. The comparison was of PGx-guided treatment selection versus TAU. Patients and raters were blinded, but PGx reports were released to treating clinicians (psychiatry, internal medicine, OB-GYN, family medicine) prior to first medication evaluation visit in the interventional arm but not in the control arm. The primary outcome was reduction in adverse drug events, and the study found no statistically significant differences between treatment groups (p=0.21). Secondary outcomes were rates of depression response (>50% reduction in 17-item HAM-D score) and remission (HAM-D score ≤7). Remission and response rates were significantly higher in the PGx-guided arm for moderate and severe depression but not for mild depression. For moderate and severe depression, the 12-week remission rate was 35% in the experimental arm and 13% in the control arm (95% CI=5–25, p=0.02). Response rates at 12 weeks were 64% vs. 46% (p=0.01). For anxiety, one secondary outcome was significant: the 12-week response rate in the PGx-guided arm was 63%, compared with 50% in the control arm (p=0.04). Patients in the PGx-guided group were more likely to have medication changes 2 weeks into the trial (81% vs. 64%) but were no more likely to have dosage adjustments. The changes in medication were in alignment with PGx report recommendations 70% of the time. Effects were not subdivided by clinician specialty. Unlike many other PGx trials, which restricted patient enrollment to those with at least one prior medication trial with inadequate response, this study had a mix of treatment-naive and treatment-resistant patients. A post hoc analysis limited to the treatment-resistant patients found similar statistically significant effects of the secondary outcomes in the pooled analysis.

Pillcheck

Papastergiou et al., 2021. This study (36) was a 6-month RCT of 213 patients with MDD. The population was majority female (~75%) and is slightly different from others in that it recruited only patients who were dissatisfied with their current treatment. The study is of note because it was pharmacist driven, rather than treating-clinician driven: Patients were randomized to have the pharmacy team give medication recommendations to their clinicians based on PGx versus "standard of care clinical guidelines" (although these were not defined). Only patients were blinded. Prescribers were not blinded and were told whether the recommendation was coming from PGx or standard care pathways. The primary outcome was improvement in score on the Patient Health Questionnaire–9, and a statistically significant time-by-group effect (p=0.03 with a mixed-effects model) was reported. Individual time points were not primarily assessed, although at 6 months there was an improvement of 36% from baseline depression severity in the PGx group, compared with 18% for the TAU group. The study did not directly examine response or remission.

Pharmacy recommendations were more likely to be accepted by prescribing clinicians if they were the result of PGx tools versus standard guidelines.

Genecept, Version 2.0

Perlis et al., 2020. This was an 8-week RCT of 304 patients who had moderate to severe depression (37), as assessed by the Structured Interview Guide for the Hamilton Depression Rating Scale (17-item version) (SIGH-D-17), and were nonresponsive to at least one prior adequate medication trial (at least 6 weeks at an adequate dosage) in the current episode, but no more than three medication trials in total. The study population was majority female (~72%). The comparison was of PGx-guided treatment selection versus TAU. Raters and patients were blinded, but prescribers were not. The study did not meet its primary endpoint, which was change in SIGH-D-17 score ($p=0.53$) and did not meet its secondary endpoints, which were response (>50% decrease in SIGH-D-17 score; $p=0.17$) or remission (SIGH-D-17 score <7; $p=0.23$). All secondary measures were also negative except frequency of response as assessed by the CGI improvement scale (CGI-I) (score ≤3), but there were no significant differences in CGI-I scores. There were also no significant differences in adverse drug effects between the two groups. Of note, the study was designed to have a 90% power to detect a SIGH-D-17 difference of 3.1 with a 23% dropout rate, but dropout was actually only ~8%, so the study had even greater power, supporting accepting its negative result as a valid assessment of the utility of PGx.

Custom Algorithms

Shan et al., 2019. This was an 8-week RCT of 80 patients with MDD (38). The study group is of note as it comprised a Chinese population and thus is among the few study populations outside of Europe or North America. It is also of note for a high dropout rate; of the 80 patients, 48 completed the trial. Finally, the design is unique in that a single clinician treated both PGx and TAU groups. The comparison entailed use of an in-house proprietary PGx tool for treatment selection versus TAU. The study did not meet its primary endpoint of difference in 17-item HAM-D score ($p=0.21$) at 8 weeks. There were no statistically significant differences in response rate, remission rate, or rate of adverse events.

Meta-Analyses

In addition to the primary studies summarized above, there have been five stand-alone meta-analyses and one meta-analysis that was integrated into the report of a smaller trial. Meta-analyses are valid when the input data are valid and when, in addition, the studies are of comparable design. Notably, any design bias appearing in a substantial proportion of included studies will lead to the finding of statistically significant group differences if the sample size is great enough. However, that difference will reflect the design bias, not the true underlying difference in effect of the intervention being tested. The meta-analyses of PGx studies have significant technical limitations, and they do not correct for the methodological limitations of the primary trials; for example, a meta-analysis of several trials that are not fully blinded does not remedy the fact that the trials were not fully blinded. All but one analysis combined trials of multiple clinical support tools, despite evidence that the tools should be considered separately. As indicated in the opening sections of this review, there is poor concordance of medication recommendations among the currently available PGx tools, and thus combining them is like combining trials of quite different tests. All found significant effects of PGx guidance but focused on different outcomes. Most discuss the potential for bias (and differential bias) in the primary studies, but these observations were not taken into account for each analysis itself.

Rosenblat et al., 2018. This was a meta-analysis (39) of four early RCTs (27, 32, 35, 40) and two open-label cohorts (28, 29) that used a range of different PGx platforms, including GeneSight, Neuropharmagen, NeuroIDgenetix, and CNSDose, tested in studies that had considerable methodological variability. The authors pooled data to calculate the relative risk of response and relative risk of remission. Using all studies, they found that PGx was associated with higher relative risks of response (relative risk=1.36, 95% CI=1.14–1.62, $p<0.001$; N=799) and remission (relative risk=1.74, 95% CI=1.09–2.77, $p=0.02$; N=735). Including the four RCTs yielded a significant difference in response (53% vs. 41%, $p=0.02$; N=590) but not remission (33% vs. 25%, $p=0.05$; N=570).

Bousman et al., 2019. This was a meta-analysis (41) of the same four early RCTs (27, 32, 35, 40) plus the GUIDED trial (19), which collectively used a range of different PGx platforms, including GeneSight, Neuropharmagen, NeuroIDgenetix, and CNSDose. The authors used a random-effects model and found a relative risk of 1.71 (95% CI=1.17–2.48, $p=0.005$) for PGx compared with TAU in achieving remission.

Brown et al., 2020. This was a meta-analysis (42) of two early open-label studies (28, 29) and two RCTs (19, 27) that used GeneSight. The authors also included a meta-analysis (43) because it contained additional data from the 2012 and 2013 trials by Hall-Flavin et al. (28, 29) that were not published in the original studies. All of these studies used the GeneSight platform. The analysis found that symptom improvement, response rate, and remission rate were statistically significantly better in the PGx group compared with the TAU group (symptom improvement difference: 10.08%, 95% CI=1.67–18.50, $p=0.019$; response: relative risk=1.40, 95% CI=1.17–1.67, $p<0.001$; remission: relative risk=1.49, 95% CI=1.17–1.89, $p=0.001$).

Ielmini et al., 2022. This was a meta-analysis (44) of a mixed cohort of GeneSight and Neuropharmagen RCTs (19, 27, 31,

32, 35). It appears also to have included both a subanalysis (21) of the GUIDED trial (19) and the GUIDED trial independent studies, and thus >25% of the sample size in this meta-analysis appears to be erroneously duplicated patient values, as the GUIDED subanalysis (21) analyzes a subset of patients (and their values) previously reported in the primary GUIDED report (19). If that is indeed the case, it would greatly bias any conclusions of the meta-analysis. The investigators found that compared with TAU, PGx-guided treatment had greater odds of being associated with treatment response, as defined by achieving a reduction >50% in baseline HAM-D score (odds ratio=1.49, 95% CI=1.29–1.73).

Brown et al., 2022. This was a meta-analysis that combined studies that used several different PGx platforms (45) (including GeneSight, Neuropharmagen, NeuroIDgenetix, and CNSDose) comprising 4,767 patients, three open-label studies (23, 28, 29), and 10 RCTs (19, 26, 27, 31, 32, 34, 35, 37, 38, 40). The study examined only remission rates, even though this was not a primary endpoint for most of these clinical trials. The authors found that patients in the PGx arm were 1.41 times more likely to achieve remission (95% CI=1.15–1.74, p<0.001), and the effect was similar when the open-label studies were excluded (relative risk=1.46, 95% CI=1.13–1.88, p<0.003). In the three open-label studies together, the difference in remission rates was not statistically significant. The effect survived leave-one-out reanalysis. The authors also performed a subanalysis with the five GeneSight trials excluded and found that the benefit for remission survived (relative risk=1.46, 95% CI=1.02–2.09, p=0.04). The authors noted the heterogeneity in the various tests but observed that CYP2C19 and CYP2D6 were genotyped in all; however, the high risk of bias from unblinded clinicians is another shared element that could influence shared effects. The authors did not perform a subanalysis removing trials with moderate or high risk of bias (i.e., a trial with high reporting bias) (35), trials with high-risk detection bias from unblind raters (32, 34), a trial with postintervention reporting bias (29), and a trial with preintervention confounding and postintervention selection bias (28).

Tiwari et al., 2022. This was a primary trial (26), but the authors also conducted a meta-analysis combining two studies (19, 27) and found a decrease of an additional 3.33% in score on the 17-item HAM-D in PGx-guided treatment versus TAU at week 8 or 10 (p=0.039) and odds ratios of 1.44 (p=0.004) for response and 1.69 (p=0.001) for remission. The prior open-label GeneSight studies (28, 29) were not included.

DISCUSSION

Since the publication of our initial position statement in 2018 (1), nearly a dozen additional clinical trials seeking to determine the utility of combinatorial PGx testing in the treatment of depression have been published, as have half a dozen meta-analyses. The main new contribution of these studies is one of numbers: several trials have included relatively large sample sizes, and >4,000 patients have now participated in PGx studies. Of these primary trials, all but three failed to meet their primary endpoints. These utilized three separate PGx tools. The PRIME Care trial (23, 24), which did meet a primary endpoint, was open-label; the second trial (31) was small (100 patients) and only the patients were blinded, and the third (36) blinded only the patients and enrolled only patients who were dissatisfied with their current treatment. Additionally, many of the more substantive methodological critiques of early trials were not addressed or remedied in these subsequent trials. None of the trials have been fully blinded, and none report on the success of blinding that was done (such as having patients guess which study arm they are in). In all but one trial, the treating clinicians were unblinded to intervention arm, and thus the risk for performance bias and attention and ancillary treatment bias is high. Although it was sometimes possible to perform subsequent analyses to investigate the presence and magnitude of performance and attention and ancillary treatment bias, such analyses have not yet been done. In all trials, the control arm was "treatment as usual," with little attention paid to how closely that treatment aligned with standards of current clinical practice. In trials reporting TAU choices, various nonstandard choices were reported (20). In most trials, the studies were industry sponsored or had significant industry support; although industry support is not in itself problematic and historically has often been integral in completing large, well-designed, definitive trials, its coexistence with the methodological concerns reviewed above augments the concern about bias. Although the recent meta-analyses reviewed here argue for a potential modest effect of PGx on treatment outcomes, they do not resolve or attempt to correct the serious methodological limitations of the underlying trials. In particular, biased, improperly blinded trials can be expected to show modest effects favoring the intervention being studied. If the number of trials grows and the outcomes are combined, those modest effects will reach statistical significance, but they may not reflect a true intervention effect. Also, because most of the meta-analyses integrate effects across different PGx tools, there is concern that these tools actually differ substantially in their treatment recommendations (5) and cannot properly be combined in an analysis. No trials have compared different PGx tools. The addition of these new data does not alter the recommendations of our 2018 report, or the advice of the FDA, that the evidence does not support the use of currently available combinatorial PGx tools for treatment selection in major depressive disorder.

While general use is not warranted, are there any, even occasional, circumstances in which commercial PGx tools should be used? Some proponents of the use of PGx have made the argument that if a clinician is not sure what antidepressant to use next, there is no harm in ordering and using PGx. This position is not supported by the evidence.

The variants chosen for use in PGx tools, and the algorithms by which they are combined, have not been shown to be predictive of clinical efficacy or side effects. To emphasize the latter point that testing to minimize side effects is not supported, only one trial included side effects as a primary endpoint (35), and it did not find a benefit to testing. Most trials did not even include side effects as secondary outcomes, and the only one that found a benefit to side effects (32) was small and had some of the highest concern for methodological bias. The use of available tools, as noted by the FDA, may lead to delay in choosing the next medication, the choice of a medication that would otherwise not be a good next choice, or inefficient use of time during a follow-up visit or health care resources, thereby harming the patient. An appropriate choice would be to follow well-documented and freely published treatment guidelines (46–48).

At a time when many practicing clinicians are hopeful for new tools to improve the challenging state of treatment of depression, it is tempting to overinterpret the significant effects in secondary analyses from these PGx trials. For example, if PGx is not helpful in part because a large fraction of patients are "incidentally" already on medications with few predicted gene-drug interactions, or are switched to medications that "incidentally" have few predicted gene-drug interactions, could PGx be helpful for the minority subfraction with predicted significant gene-drug interactions? The subanalysis from the GUIDED trial (19) to that end is intriguing, for example. History and best practices of clinical trials emphasize and reemphasize that trials should be designed and powered to test a primary endpoint, and that any secondary endpoints are very useful in guiding future research but in general should not be applied to guiding practice, as they are less likely to be true intervention effects, especially when there are not rigorous design controls and other indirect evidence to support an effect (49–51). In other words, with each new hypothesis that is tested in a subgroup or post hoc analysis, the risk of false positive effects or type I error is markedly elevated. Thus, the heavy reliance on post hoc analyses in the existing PGx literature is enriched for the potential of this type of distortion, and the post hoc findings need to be tested in rigorous follow-up trials to determine whether they are real intervention effects. Post hoc analyses are designed not to inform clinical practice but to prompt further investigation.

Moreover, some (30) have raised concerns that the use of PGx tools may bias clinicians toward less conventional treatment choices—for example, newer medications such as desvenlafaxine (which has no known gene-drug interactions because it is newer, and is not metabolized by CYP450 enzymes) or older medications (which have fewer studies on how the drug interacts with genetic variants, as pointed out by McCarthy et al. [34]). Clinicians prescribing antidepressants for adolescents with MDD in the PGx trials were statistically less likely to use SSRIs in the PGx-guided arm despite SSRIs' being the only FDA-approved agents for treatment of MDD in adolescents. The development, translation, and adoption of new approaches and technologies should be evidence based. The onus of proof for change has not been met by combinatorial PGx in its current state.

Although beyond the scope of this review, an important part of future work will be to understand the factors that make the current state of the treatment of depression so challenging and the hope for help from new tools like pharmacogenomics so visceral—for example, clinical heterogeneity, practical limits in access to some modalities of treatment (e.g., specialized psychotherapy), and social structures. However, use of new tools should always be evidence based.

To that end, we highlight again that fully blinded prospective studies are crucial in resolving the question of whether PGx is useful, now or in future versions.

Fully Blinded PGx Studies: The Most Important Next Step

The requirement of blinding in clinical drug trials is well known, and the same concerns apply to PGx trials. It is important to emphasize that receiving genetic information alone can result in placebo or nocebo effects that can bias responses (and perhaps even physiology) and that this effect is seen even when the genetic information disclosed is false or irrelevant (e.g., 52). Moreover, there is increasing evidence that there is a "seductive allure" of any high-tech or scientific information, whether or not it is relevant to the intervention being proposed. For example, including irrelevant neuroscientific information increases how satisfying participants rate an explanation of a psychological phenomenon (summarized in reference 53). The fact that one of the trials we reviewed (36) found that clinicians were more likely to accept recommendations on treatment provided by pharmacists when that recommendation was based on PGx than when it was based on gold-standard clinical pharmacological practice reinforces the concern that unblinded treating clinicians may be influenced by receiving PGx information. Thus, our recommendation is to blind treating clinicians, patients, and raters and incorporate a metric to determine how successful that blinding was for each party.

It is interesting to observe that the feasibility of a fully blinded PGx study is disputed by most of the authors of the primary studies that were reviewed here, often on ethical grounds. The argument is generally structured as follows: 1) to fully blind the study, clinicians would need to be given falsified genotyping information, and 2) providing falsified genotyping information is unethical. Point one of this argument, we believe, is easily challenged. Given that PGx reports provide recommendations on particular medications, it would be possible, even easy, to provide a report of recommendations without providing clinicians any genotyping information; indeed, a trial we reviewed (34) was able to provide a sham PGx report (albeit an imperfect sham) with approval of the institutional review board, and we are pleased to report that an ongoing trial (Australian and New Zealand Clinical Trial Registry identifier ACTRN12621000181808)

will provide blinded clinicians identically formatted reports based on either PGx or gold-standard clinical practice guidelines, as we have previously recommended. Point two of the argument is more ethically complex but also not true prima facie. For example, if the possibility of false genotypes on a report is disclosed during trial enrollment; if it is agreed not to use genotype information outside the bounds of the trial; if it is understood that true genotypes will be disclosed at the end of the trial; and if it is discussed that there is great uncertainty about whether the genotypic information matters for clinical decision making in the treatment of depression (and thus being given false genotyping information is in equipoise with being given true genotyping information), then such a trial design may indeed be permissible. It has been pointed out that the most relevant comparison would be between standard protocol-based treatment and PGx-guided treatment, and such a study could be designed and implemented. Such studies are the standard for medication trials, so they are accepted, even required, ethical practices. Moreover, such an approach of transparency may in fact be morally preferable to some previous approaches, where unblinded treating clinicians were asked to knowingly dissemble to patients in a TAU arm to maintain the patient blind by discussing a mock report with names of "sham drugs" (19; see the article's supplement).

There are several additional improvements to study design that may be considered, including, for example, more systematic accounting of the estimated 60% of the variance in drug response that is not due to common genetic variation (7). However, we emphasize that fully blinded PGx studies are the most important next step.

Resetting PGx Expectations in Psychiatry by Examining Other Specialties

How high should our expectations be for the usefulness of PGx for treatment guidance in depression? In a qualitative study (54) on clinician attitudes toward PGx leading up to the PRIME Care trial, the authors reported both hope and skepticism about the potential for PGx to improve treatment selection in depression. Some of the most thoughtful skepticism came from the primary care clinicians, who have seen how the hope of personalized care through genetics has played out so far in other areas of medicine. They pointed out the experience with warfarin treatment, where the influence of genetic variants has been best defined: specifically, despite the ability to obtain well-validated and relevant genetic information, this information has not changed clinical practice. For example, a variant at the *VKORC1* locus accounts for 25% of the phenotypic variability in warfarin dosing. Yet, meta-analyses of RCTs do not show a benefit to gene-guided therapy compared to treatment as usual, and the consensus is that testing for these variants is not recommended (55, 56).

Another example is that of clopidogrel. CYP2C19 is necessary to metabolize clopidogrel into an active compound (clopidogrel is a prodrug). Despite promising effects of genotyping in guiding the selection of this antiplatelet drug in early observational trials, follow-up with well-powered RCTs did not find a benefit to PGx-guided treatment, even when restricting the study to patients with loss-of-function (LoF) alleles. For example, in the TAILOR-PCI study, 903 patients with LoF alleles who were in the genotype-guided arm and received ticagelor did not have significantly different cardiovascular endpoints than 945 patients with LoF alleles who were in the control arm and received clopidogrel (17). An additional well-powered RCT, the Popular Genetics Trial, similarly did not show a benefit from genotype-guided medication selection (18). Although not without controversy, the routine testing of patients for clopidogrel selection is not generally recommended (57–59).

As for the relevance of this to the treatment of depression, these examples should make it clear that even when a genetic variant has a clear effect on pharmacokinetics, further careful trials are required before it can be decided whether that effect makes a difference to patient outcomes. In addition, no gene or combination of genes used in PGx tests have yet been shown to approach the predictive value of genetic testing in other branches of medicine. This may be because so many genes contribute to risk and response, with none determining more than modest proportions of those outcomes. Alternatively, there may be genes that are good predictors that have not yet been identified.

In fact, the place where pharmacogenetic information has been most useful in general medicine has not been in improving benefits to patients, but in estimating risk of rare catastrophic adverse events, a goal that is perhaps underpursued in psychiatry. The prime example is testing for HLA-B*57:01 prior to initiating abacavir for the treatment of HIV infection. Although the frequency of this allele is in the range of <1% to ~8%, depending on ancestry, those who possess it have a 55% chance of developing a severe, CD8-T-cell–mediated drug hypersensitivity reaction that is otherwise rare across the general population (60). Testing for HLA-B*57:01 is now considered a standard of care (61). Investigating risk of severe adverse drug reactions is far from the current focus of PGx research in psychiatry, although there is potential for clinical support tools to be helpful. For example, variants in HLA have been found to influence risk of Stevens-Johnson syndrome with initiation of carbamazepine and oxcarbazepine (62) and to influence the risk of agranulocytosis with initiation of clozapine (63, 64). To our knowledge, it is uncertain how useful PGx tools might be for rare adverse drug events, such as severe hyponatremia from medication-induced syndrome of inappropriate antidiuretic hormone secretion. Only one of the PGx studies we reviewed here had adverse drug events as a primary outcome (with no significant effect found); many of the trials detected no effect on side effects, or did not reports on side effects.

These modest clinical practice changes achievable from genetic information on how the body processes and responds to common medications are in contrast with advances in cancer therapeutics. Based on genotyping driver mutations

in tumors, the field of oncology has seen the co-development of specialized drugs alongside genetic biomarkers, such as BRAF V600E targeted treatment (65). Such co-development of novel medications with targeted genetic information in psychiatry would represent a large shift in approach to the way PGx research and its applications are currently being pursued in oncology.

Clearing the Path to Future Progress by Addressing Common Misconceptions About PGx in the Treatment of Depression

We conclude this piece by discussing four myths about PGx clinical support tools in the treatment of depression.

1. Myth 1: Because a person's genes do not change over time, PGx clinical support tools represent a onetime cost, and the information can be used in the future or in other clinical contexts.

Although it is true that a person's structural genome is largely invariant, having the results of one PGx test in the past is unlikely to be equivalent to tests in the future. As we reviewed in the introduction, there is considerable variation as to which loci current tests examine, which alleles at each locus they examine, and the proprietary (and thus opaque) algorithm that is used to combine the effects of each allele for a given individual. If a new locus or allele is found to be important in the future, a current test would not have included it in the testing panel. Similarly, even if a company maintains the same panel of variants but changes its proprietary algorithm in the future, the patient's prior results would need to be reanalyzed. While the specific evaluation of the utility of these individual tests is beyond this review, which considers combinatorial tools, we did point out in the introduction that the FDA table lists four instances where CYP2D6 poor-metabolizer status might be relevant to four medications used for depression; if genotype is reported in a PGx report, it might be able to provide such a genotype as "raw data" that is separate from the combinatorial algorithm, but to do so would require careful examination of the report for inclusion of all potentially relevant genotypes at the genetic locus and should not be assumed.

In addition, while gene structure changes little in a lifetime, gene expression and the activity of gene products can change dramatically. This is well illustrated for the primary elements of commercial tools marketed for choosing antidepressant medications. Cytochrome enzymes that metabolize medications show variation in expression and activity based on diet, exercise, and concomitant drug use, including the use of over-the-counter and recreational drugs. These effects are generally much larger than those seen among different genetic (structural) variants coding for these enzymes (66).

2. Myth 2: Because genetic variants account for a large part of variation in treatment outcomes in depression, PGx tools are useful prima facie.

It is often cited that common genetic variants may account for up to 40% of variation in antidepressant response (e.g., 7). Although this does point to the potential of predicting response through investigating the genome, it does not mean that existing PGx tools (especially the products currently offered) should be useful in choosing pharmacotherapy. First, current tools examine only a fraction of the genetic loci thought to be involved in antidepressant response. Second, even if a tool were to accurately capture the variants responsible for the entire genetic contribution to antidepressant response, it would not necessarily be true that using a different antidepressant would lead to improved antidepressant response (e.g., if the physiology that flows from those variants would yield preferentially to a nontypical pharmacological approach). The fact that our diagnostic categories are defined by symptoms and likely represent a heterogeneity of underlying pathophysiological abnormalities further complicates the use of genetics tools. The usefulness of PGx tools must therefore be rigorously and empirically tested for each particular use.

3. Myth 3: Broad use of PGx clinical support tools will reduce bias and inequities in outcomes.

We have noted recent claims in the literature (e.g., 67) that unequal access to PGx could augment existing inequities in treatment outcomes in depression, and that the solution would be to ensure broad implementation of PGx and removal of financial or geographic barriers to access to combinatorial PGx tools. Another related example is the recent campaign from Myriad's "Building Trust by Reducing Bias—GeneSight as a Tool for Mental Health Equity" (see https://genesight.com/genesight-cares/). Such concern is only warranted if PGx improves outcomes for the population of interest, and it is not at all clear that PGx testing improves outcomes in any group of patients. Moreover, it is an open question as to how different genetic ancestries affect PGx utility; most of the studies to date likely have overrepresentation of European genetic ancestries, rendering sample sizes too small to examine other ancestries. Similarly, it is unclear to what extent the current tests incorporate or are affected by genetic alleles present at different frequencies in different genetic ancestries or by untested genes that differ among people of various ancestries. Studies generally have not explicitly commented on the prevalence of different genetic ancestries or how this might affect the distribution of the common variants present in their analyses. These are potential future projects, and potentially fruitful ones, and we are encouraged that such projects have become more of a focus for the academic and biotech research communities, but offering an unproven product to more people is not an appropriate direction of care or a good use of limited resources.

4. Myth 4: It is prudent from a defensive medicine perspective to use PGx clinical support tools prior to

prescribing medications, in case a patient experiences an adverse event.

We have noticed a claim from some that there could be adverse legal consequences to prescribing in ignorance of genomic metabolizer information. Forensic use of metabolizer genetic testing has been used to reduce liability of prescribers—for example, as part of arguments for why codeine could have led to overdose in a child without the physician prescribing a recklessly high dose—but there is also concern that "lawsuits might try to insinuate that doctors of patients who suffered adverse reactions were negligent in not performing pharmacogenetic tests beforehand" (68). Such cases would usually be decided on the basis of what is determined to be "standard of care." Standard of care is a complicated legal construct affected by current guidelines from specialty societies. For example, testing would be standard of care prior to starting abacavir, but not warfarin, even though the genetic variant does affect warfarin blood levels and is on the drug's FDA label. These cases may also serve as guide points for clinicians in how to handle medications such as citalopram, for which the FDA table and label list higher risk of QTc prolongation for CYP2D6 poor metabolizers. QTc-prolonging medications have existing guidelines on how to monitor and mitigate cardiac risks, which do not depend on PGx testing. With the prior 2018 position from our workgroup, and the FDA's caution against decision making based on PGx testing, a prescriber should plausibly be on firmer defensive ground by avoiding making treatment decisions based on PGx testing than by using it. Of course, the legal landscape is complex and shifting, and our discussion should not be considered legal advice.

CONCLUSIONS

We have reviewed the new evidence on the potential usefulness of combinatorial PGx testing in the treatment of depression, focusing on studies published since our last position piece. The addition of these new data does not alter the recommendations of our 2018 report, or the advice of the FDA, that the evidence does not support the use of currently available combinatorial PGx tools for treatment selection in major depressive disorder. Genetic approaches remain promising, and we look forward to future studies and advances in the field. However, we advise devoting greater attention to implementing study designs consistent with other studies of treatment interventions. This includes appropriate blinding and assessment of the success of blinding. We suggest that PGx tests be based firmly on data regarding the relationship of gene variants to response. Improved statistical analysis, including correction for multiple measures, is also required.

AUTHOR AND ARTICLE INFORMATION

Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston (Baum); Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minneapolis (Widge); Butler Hospital Neuromodulation Research Facility, Providence, R.I., and Department of Psychiatry and Human Behavior, Alpert Medical School, Brown University, Providence, R.I. (Carpenter); Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta (McDonald); Department of Psychiatry and Program for Neuropsychiatric Research, McLean Hospital, Harvard Medical School, Belmont, Mass. (Cohen); Department of Psychiatry and Behavioral Sciences, Dell Medical School, University of Texas, Austin (Nemeroff).

Send correspondence to Dr. Nemeroff (cnemeroff@austin.utexas.edu).

Dr. Baum receives royalties from Oxford University Press. Dr. Widge has received donations of deep brain stimulation devices from Medtronic and consulting fees from Abbott, and he holds multiple granted and pending patents related to deep brain stimulation. Dr. Carpenter has received research support from Affect Neuro, Janssen Pharmaceuticals, Neuronetics, Neurolief, Nexstim, and SynapseBio, and she has served as a scientific adviser or consultant for Affect Neuro, Janssen Pharmaceuticals, Neuronetics, Neurolief, Nexstim, Otsuka, Sage Therapeutics, and Sunovion. Dr. McDonald is a Deputy Editor of the *American Journal of Psychiatry*; disclosures of Editors' financial relationships appear in the April 2024 issue of the *Journal*. Dr. Nemeroff has served as a consultant for AbbVie, ANeuroTech, Autobahn Therapeutics, BioXcel Therapeutics, Clexio, EMA Wellness, EmbarkNeuro, Engrail Therapeutics, GoodCap Pharmaceuticals, Guidepoint, Intra-Cellular Therapies, Magstim, Ninnion Therapeutics, Pasithea Therapeutics Corp., Relmada Therapeutics, Sage, Senseye, Signant Health, Silo Pharma, and SynapseBio; he is a stockholder in Antares, Corcept Therapeutics, EMA Wellness, Naki Health, Precise-Ment Health, Relmada Therapeutics, and Seattle Genetics; he has served on scientific advisory boards for ANeuroTech, the Anxiety and Depression Association of America (ADAA), the Brain and Behavior Research Foundation, Galen Health, Heading Health, the Laureate Institute for Brain Research, Pasithea Therapeutics Corp., Sage, Signant Health, and Skyland Trail and on the boards of directors for ADAA, Gratitude America, and Lucy Scientific Discovery; and he is named on patents on a method and devices for transdermal delivery of lithium and a method of assessing antidepressant drug therapy via transport inhibition of monoamine neurotransmitters by ex vivo assay. Dr. Cohen reports no financial relationships with commercial interests.

Received August 21, 2023; revisions received October 30, December 5, 2023, and January 4, 2024; accepted January 17, 2024; published online April 30, 2024.

REFERENCES

1. Zeier Z, Carpenter LL, Kalin NH, et al: Clinical implementation of pharmacogenetic decision support tools for antidepressant drug prescribing. Am J Psychiatry 2018; 175:873–886
2. FDA issues warning letter to genomics lab for illegally marketing genetic test that claims to predict patients' responses to specific medications [press release]. 2019. https://www.fda.gov/news-events/press-announcements/fda-issues-warning-letter-genomics-lab-illegally-marketing-genetic-test-claims-predict-patients
3. International Society of Psychiatric Genetics: Genetic Testing and Psychiatric Disorders: A Statement From the International Society of Psychiatric Genetics. 2019. https://ispg.net/genetic-testing-statement/
4. Smith TL, Nemeroff CB: Pharmacogenomic testing and antidepressant response: problems and promises. Braz J Psychiatry 2020; 42:116–117
5. Bousman CA, Dunlop BW: Genotype, phenotype, and medication recommendation agreement among commercial pharmacogenetic-based decision support tools. Pharmacogenomics J 2018; 18:613–622
6. Rush AJ, Trivedi MH, Wisniewski SR, et al: Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 2006; 163:1905–1917
7. Tansey KE, Guipponi M, Hu X, et al: Contribution of common genetic variants to antidepressant response. Biol Psychiatry 2013; 73:679–682

8. Yoshida K, Takahashi H, Higuchi H, et al: Prediction of antidepressant response to milnacipran by norepinephrine transporter gene polymorphisms. Am J Psychiatry 2004; 161:1575–1580
9. Marshe VS, Maciukiewicz M, Rej S, et al: Norepinephrine transporter gene variants and remission from depression with venlafaxine treatment in older adults. Am J Psychiatry 2017; 174: 468–475
10. Chappell K, Ait Tayeb AEK, Colle R, et al: The association of ARRB1 polymorphisms with response to antidepressant treatment in depressed patients. Front Pharmacol 2022; 13:974570
11. Binder EB, Nemeroff CB: The CRF system, stress, depression and anxiety: insights from human genetic studies. Mol Psychiatry 2010; 15:574–588
12. Rosenblat JD, Goldberg JF, McIntyre RS: Consumer warning for genetic tests claiming to predict response to medications: implications for psychiatry [letter]. Am J Psychiatry 2019; 176:412–413
13. Jeffrey Shuren, MD, JD, director of the FDA's Center for Devices and Radiological Health and Janet Woodcock, MD, director of the FDA's Center for Drug Evaluation and Research on agency's warning to consumers about genetic tests that claim to predict patients' responses to specific medications [press release]. US Food and Drug Administration, November 1,2018. https://www.fda.gov/news-events/press-announcements/jeffrey-shuren-md-jd-director-fdas-center-devices-and-radiological-health-and-janet-woodcock-md
14. FDA announces collaborative review of scientific evidence to support associations between genetic information and specific medications [press release]. US Food and Drug Administration, February 20, 2020. https://www.fda.gov/news-events/press-announcements/fda-announces-collaborative-review-scientific-evidence-support-associations-between-genetic
15. US Food and Drug Administration: Table of Pharmacogenetic Associations. 2022. https://www.fda.gov/medical-devices/precision-medicine/table-pharmacogenetic-associations
16. Clinical Pharmacogenetics Implementation Consortium: Guidelines. https://cpicpgx.org/guidelines/
17. Pereira NL, Farkouh ME, So D, et al: Effect of genotype-guided oral P2Y12 inhibitor selection vs conventional clopidogrel therapy on ischemic outcomes after percutaneous coronary intervention: the TAILOR-PCI randomized clinical trial. JAMA 2020; 324: 761–771
18. Claassens DMF, Vos GJA, Bergmeijer TO, et al: A genotype-guided strategy for oral P2Y(12) inhibitors in primary PCI. N Engl J Med 2019; 381:1621–1631
19. Greden JF, Parikh SV, Rothschild AJ, et al: 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 2019; 111:59–67
20. Zubenko GS, Sommer BR, Cohen BM: On the marketing and use of pharmacogenetic tests for psychiatric treatment. JAMA Psychiatry 2018; 75:769–770
21. Thase ME, Parikh SV, Rothschild AJ, et al: Impact of pharmacogenomics on clinical outcomes for patients taking medications with gene-drug interactions in a randomized controlled trial. J Clin Psychiatry 2019; 80:19m12910
22. Forester BP, Parikh SV, Weisenbach S, et al: Combinatorial pharmacogenomic testing improves outcomes for older adults with depression. Am J Geriatr Psychiatry 2020; 28:933–945
23. Oslin DW, Lynch KG, Shih MC, et al: 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 2022; 328:151–161
24. Oslin DW, Chapman S, Duvall SL, et al: Study design and implementation of the Precision Medicine in Mental Health Care (PRIME Care) trial. Contemp Clin Trials 2021; 101:106247
25. Ramsey CM, Lynch KG, Thase ME, et al: Prevalence of predicted gene-drug interactions for antidepressants in the treatment of major depressive disorder in the Precision Medicine in Mental Health Care Study. J Affect Disord 2021; 282:1272–1277
26. Tiwari AK, Zai CC, Altar CA, et al: Clinical utility of combinatorial pharmacogenomic testing in depression: a Canadian patient- and rater-blinded, randomized, controlled trial. Transl Psychiatry 2022; 12:101
27. Winner JG, Carhart JM, Altar CA, et al: A prospective, randomized, double-blind study assessing the clinical impact of integrated pharmacogenomic testing for major depressive disorder. Discov Med 2013; 16:219–227
28. Hall-Flavin DK, Winner JG, Allen JD, et al: Utility of integrated pharmacogenomic testing to support the treatment of major depressive disorder in a psychiatric outpatient setting. Pharmacogenet Genomics 2013; 23:535–548
29. Hall-Flavin DK, Winner JG, Allen JD, et al: Using a pharmacogenomic algorithm to guide the treatment of depression. Transl Psychiatry 2012; 2:e172
30. Vande Voort JL, Orth SS, Shekunov J, et al: A randomized controlled trial of combinatorial pharmacogenetics testing in adolescent depression. J Am Acad Child Adolesc Psychiatry 2022; 61: 46–55
31. Han C, Wang SM, Bahk WM, et al: A pharmacogenomic-based antidepressant treatment for patients with major depressive disorder: results from an 8-week, randomized, single-blinded clinical trial. Clin Psychopharmacol Neurosci 2018; 16:469–480
32. Pérez V, Salavert A, Espadaler J, et al: Efficacy of prospective pharmacogenetic testing in the treatment of major depressive disorder: results of a randomized, double-blind clinical trial. BMC Psychiatry 2017; 17:250
33. Menchón JM, Espadaler J, Tuson M, et al: Patient characteristics driving clinical utility in psychiatric pharmacogenetics: a reanalysis from the AB-GEN multicentric trial. J Neural Transm (Vienna) 2019; 126:95–99
34. McCarthy MJ, Chen Y, Demodena A, et al: A prospective study to determine the clinical utility of pharmacogenetic testing of veterans with treatment-resistant depression. J Psychopharmacol 2021; 35: 992–1002
35. Bradley P, Shiekh M, Mehra V, et al: Improved efficacy with targeted pharmacogenetic-guided treatment of patients with depression and anxiety: a randomized clinical trial demonstrating clinical utility. J Psychiatr Res 2018; 96:100–107
36. Papastergiou J, Quilty LC, Li W, et al: Pharmacogenomics guided versus standard antidepressant treatment in a community pharmacy setting: a randomized controlled trial. Clin Transl Sci 2021; 14: 1359–1368
37. Perlis RH, Dowd D, Fava M, et al: Randomized, controlled, participant- and rater-blind trial of pharmacogenomic test-guided treatment versus treatment as usual for major depressive disorder. Depress Anxiety 2020; 37:834–841
38. Shan X, Zhao W, Qiu Y, et al: Preliminary clinical investigation of combinatorial pharmacogenomic testing for the optimized treatment of depression: a randomized single-blind study. Front Neurosci 2019; 13:960
39. Rosenblat JD, Lee Y, McIntyre RS: The effect of pharmacogenomic testing on response and remission rates in the acute treatment of major depressive disorder: a meta-analysis. J Affect Disord 2018; 241:484–491
40. Singh AB: Improved antidepressant remission in major depression via a pharmacokinetic pathway polygene pharmacogenetic report. Clin Psychopharmacol Neurosci 2015; 13:150–156
41. Bousman CA, Arandjelovic K, Mancuso SG, et al: Pharmacogenetic tests and depressive symptom remission: a meta-analysis of randomized controlled trials. Pharmacogenomics 2019; 20:37–47
42. Brown L, Vranjkovic O, Li J, et al: The clinical utility of combinatorial pharmacogenomic testing for patients with depression: a meta-analysis. Pharmacogenomics 2020; 21:559–569
43. Altar CA, Carhart J, Allen JD, et al: Clinical utility of combinatorial pharmacogenomics-guided antidepressant therapy: evidence

from three clinical studies. Mol Neuropsychiatry 2015; 1: 145–155
44. Ielmini M, Caselli I, Critelli F, et al: Efficacy and tolerability of therapies set under pharmacogenetic tools suggestions: a systematic review with meta-analysis about mood disorders. Psychiatry Res 2022; 311:114482
45. Brown LC, Stanton JD, Bharthi K, et al: Pharmacogenomic testing and depressive symptom remission: a systematic review and meta-analysis of prospective, controlled clinical trials. Clin Pharmacol Ther 2022; 112:1303–1317
46. Cleare A, Pariante CM, Young AH, et al: Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2008 British Association for Psychopharmacology guidelines. J Psychopharmacol 2015; 29:459–525
47. Kennedy SH, Lam RW, McIntyre RS, et al: Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: section 3. Pharmacological treatments. Can J Psychiatry 2016; 61:540–560
48. McQuaid JR, Buelt A, Capaldi V, et al: The management of major depressive disorder: synopsis of the 2022 US Department of Veterans Affairs and US Department of Defense Clinical Practice Guideline. Ann Intern Med 2022; 175:1440–1451
49. Stanley K: Design of randomized controlled trials. Circulation 2007; 115:1164–1169
50. Andrade C: The primary outcome measure and its importance in clinical trials. J Clin Psychiatry 2015; 76:e1320–e1323
51. Freemantle N: Interpreting the results of secondary end points and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic? BMJ 2001; 322:989–991
52. Turnwald BP, Goyer JP, Boles DZ, et al: Learning one's genetic risk changes physiology independent of actual genetic risk. Nat Hum Behav 2019; 3:48–56
53. Aono D, Yaffe G, Kober H: Neuroscientific evidence in the courtroom: a review. Cogn Res Princ Implic 2019; 4:40
54. Vest BM, Wray LO, Brady LA, et al: Primary care and mental health providers' perceptions of implementation of pharmacogenetics testing for depression prescribing. BMC Psychiatry 2020; 20:518
55. Stergiopoulos K, Brown DL: Genotype-guided vs clinical dosing of warfarin and its analogues: meta-analysis of randomized clinical trials. JAMA Intern Med 2014; 174:1330–1338
56. Belley-Cote EP, Hanif H, D'Aragon F, et al: Genotype-guided versus standard vitamin K antagonist dosing algorithms in patients initiating anticoagulation: a systematic review and meta-analysis. Thromb Haemost 2015; 114:768–777
57. Tantry US, Hennekens CH, Zehnder JL, et al: Clopidogrel resistance and clopidogrel treatment failure. Waltham, Mass, UpToDate. https://www.uptodate.com/contents/clopidogrel-resistance-and-clopidogrel-treatment-failure
58. Nissen SE: Pharmacogenomics and clopidogrel: irrational exuberance? JAMA 2011; 306:2727–2728
59. Bauer T, Bouman HJ, van Werkum JW, et al: Impact of CYP2C19 variant genotypes on clinical efficacy of antiplatelet treatment with clopidogrel: systematic review and meta-analysis. BMJ 2011; 343:d4588
60. Mallal S, Phillips E, Carosi G, et al: HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med 2008; 358:568–579
61. Saag MS, Gandhi RT, Hoy JF, et al: Antiretroviral drugs for treatment and prevention of HIV infection in adults: 2020 recommendations of the International Antiviral Society–USA panel. JAMA 2020; 324:1651–1669
62. Tangamornsuksan W, Chaiyakunapruk N, Somkrua R, et al: Relationship between the HLA-B*1502 allele and carbamazepine-induced Stevens-Johnson syndrome and toxic epidermal necrolysis: a systematic review and meta-analysis. JAMA Dermatol 2013; 149:1025–1032
63. Goldstein JI, Jarskog LF, Hilliard C, et al: Clozapine-induced agranulocytosis is associated with rare HLA-DQB1 and HLA-B alleles. Nat Commun 2014; 5:4757
64. Girardin FR, Poncet A, Perrier A, et al: Cost-effectiveness of HLA-DQB1/HLA-B pharmacogenetic-guided treatment and blood monitoring in US patients taking clozapine. Pharmacogenomics J 2019; 19: 211–218
65. Bollag G, Tsai J, Zhang J, et al: Vemurafenib: the first drug approved for BRAF-mutant cancer. Nat Rev Drug Discov 2012; 11:873–886
66. Zanger UM, Schwab M: Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol Ther 2013; 138:103–141
67. Leitch TM, Killam SR, Brown KE, et al: Ensuring equity: pharmacogenetic implementation in rural and tribal communities. Front Pharmacol 2022; 13:953142
68. Hutson S: Pharmacogenetics raises new legal questions. Nat Med 2010; 16:729
69. Dubovsky SL: The limitations of genetic testing in psychiatry. Psychother Psychosom 2016; 85:129–135
70. Rosenblat JD, Lee Y, McIntyre RS: Does pharmacogenomic testing improve clinical outcomes for major depressive disorder? A systematic review of clinical trials and cost-effectiveness studies. J Clin Psychiatry 2017; 78:720–729
71. Zhao M, Ma J, Li M, et al: Cytochrome P450 enzymes and drug metabolism in humans. Int J Mol Sci 2021; 22:12808
72. Konstandi M, Johnson EO: Age-related modifications in CYP-dependent drug metabolism: role of stress. Front Endocrinol (Lausanne) 2023; 14:1143835
73. Macaluso M, Preskorn SH: Knowledge of the pharmacology of antidepressants and antipsychotics yields results comparable with pharmacogenetic testing. J Psychiatr Pract 2018; 24:416–419

Continuing Medical Education

You can earn CME credits by reading this article. Three articles in every *American Journal of Psychiatry* issue comprise a short course for up to 1 *AMA PRA Category 1 Credit*™ each. The course consists of reading the article and answering three multiple-choice questions with a single correct answer. CME credit is issued only online. Readers who want credit must subscribe to the AJP Continuing Medical Education Course Program (psychiatryonline.org/cme), select *The American Journal of Psychiatry* at that site, take the course(s) of their choosing, complete an evaluation form, and submit their answers for CME credit. A certificate for each course will be generated upon successful completion. This activity is sponsored by the American Psychiatric Association.

Examination Questions for "Pharmacogenomic Clinical Support Tools for the Treatment of Depression"

1. **Which use of commercial combinatorial PGx clinical support tools is supported by the current evidence in regard to the treatment of depression?**
   - A. When there has been inadequate response to at least one antidepressant medication, in order to select an antidepressant that is most likely to produce the greatest magnitude of response
   - B. The evidence does not support the use of currently available combinatorial PGx tools for treatment selection in major depressive disorder
   - C. When there has been inadequate response to at least one antidepressant medication, in order to select an antidepressant that is most likely to produce the greatest speed of response
   - D. When there has been inadequate response to at least one antidepressant medication, in order to select an antidepressant that is most likely to be tolerated without side effects
2. **What best describes the common clinical study design that has been missing from existing trials of PGx clinical support tools for depression, and is the most important next step in PGx research?**
   - A. Fully blinded study design, where patients, clinicians, and raters are all blinded to PGx arm
   - B. Multi-center randomized controlled trial
   - C. Open-label trial
   - D. A double blinded trial, where patients and raters are blinded to PGx arm, and clinicians use the true PGx report
3. **Which of the following are true about current pharmacogenomic clinical support tools in the treatment of depression?**
   - A. Because a person's genes do not change over time, PGx clinical support tools represent a onetime cost, and the information can be used in the future or in other clinical contexts
   - B. Because genetic variants account for a large part of variation in treatment outcomes in depression, PGx tools are useful prima facie
   - C. It is prudent from a defensive medicine perspective to use PGx clinical support tools prior to prescribing medications, in case a patient experiences an adverse event
   - D. None of these are true; all of the above are best thought of as myths about PGx clinical support tools