Received: 24 October 2019 | Revised: 21 February 2020 | Accepted: 27 March 2020

DOI: 10.1002/da.23029

RESEARCH ARTICLE

ANXIETY AND DEPRESSION ASSOCIATION OF AMERICA WILEY

Randomized, controlled, participant- and rater-blind trial of pharmacogenomic test-guided treatment versus treatment as usual for major depressive disorder

Roy H. Perlis[1] | Daniel Dowd[2] | Maurizio Fava[1] | Todd Lencz[3] | David S. Krause[2]

[1]Department of Psychiatry and Division of Clinical Research, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts

[2]Clinical Research & Development and Medical Affairs, Genomind, Inc., King of Prussia, Pennsylvania

[3]Department of Psychiatry, Zucker School of Medicine at Hofstra/Northwell, Hempstead, New York

Correspondence
David S. Krause, Clinical Research & Development and Medical Affairs, Genomind, Inc., 2200 Renaissance Boulevard, King of Prussia, PA 19406.
Email: dkrause@genomind.com

Funding information
Genomind, Inc.

Abstract

Background: Cohort and cost-effectiveness studies suggest that measuring variation in genes that influence metabolism of common drugs could improve antidepressant treatment outcomes. Prior randomized trials have yielded inconsistent results.

Method: Multicenter randomized double-blind (subject and rater), controlled trial of pharmacogenomic testing among outpatients with nonpsychotic major depressive disorder. Study participants ($n = 304$) were randomized 1:1 to assay-guided treatment (AGT; $N = 151$) or treatment-as-usual (TAU; $N = 153$). Participants and raters were blinded to study arm; unblinded clinicians received results of a pharmacogenomic test and adjusted treatment in light of the test report. Primary outcome was change over 8 weeks in Hamilton Depression Rating Scale (SIGH-D-17).

Results: For the primary comparison of interest, change in SIGH-D-17, no significant difference was detected between AGT and TAU at Week 8 ($p = .53$). Rates of study completion also did not differ between the arms (AGT 92.7%, TAU 92.2% ($\chi^2 = 0.03$, $df = 1$, $p = .86$). Exploratory analyses suggested significantly fewer individuals experienced worsening of depressive symptoms following AGT, and that treatment concordant with assay results was associated with greater likelihood of remission.

Conclusion: Pharmacogenomic testing using a panel of pharmacokinetic and pharmacodynamic variants was not associated with significant improvement in the primary efficacy outcome when providers were unconstrained by the assay results. Further investigation is needed to understand the discordance with cost-effectiveness results and among randomized trials.

KEYWORDS

antidepressants, clinical trials, depression, genetics, pharmacotherapy

1 | INTRODUCTION

More than two decades of study indicates that functional variations in a small number of pharmacokinetic genes influence blood levels, which may in turn influence treatment response, for a majority of the pharmacopeia. This includes more than 80% of medications commonly used in psychiatric treatment (Cacabelos, Cacabelos, & Carril, 2019), and is reflected in Food and Drug Administration labels for at least 36 medications commonly used in psychiatry or neurology (Food and Drug Administration, 2018).

Cohort studies using claims data or electronic health records found that treatment with medications metabolized through the cytochrome P450 system was associated with greater medical cost and readmission, after accounting for differences in comorbidity, suggesting an opportunity to reduce cost by measuring this variation (McCoy, Castro, Cagan, Roberson, & Perlis, 2017). Naturalistic cost-

effectiveness studies likewise suggest that testing for common genetic variation may be associated with improved treatment outcomes. For example, two prior studies using a commercial test found economically meaningful and statistically significant reduction in health care utilization over 4 or 6 months (Fagerness, Fonseca, & Hess, 2014; Perlis, Mehta, Edwards, Tiwari, & Imbens, 2018).

However, efforts to demonstrate that pharmacogenomic testing improves efficacy in short-term treatment using randomized, controlled trials have yielded mixed results. Two single-blind 12-week studies have identified significant short-term benefit on their primary outcome measure (Bradley, Shiekh, & Mehra, 2018; Singh, 2015) Conversely, two single-blind studies failed to distinguish significant benefit on primary outcome measures (Pérez, Salavert, & Espadaler, 2017), including a recent study enrolling more than 1,100 participants (Greden, Parikh, & Rothschild, 2019). This apparent discordance suggests the importance of continued investigation to understand the potential benefit, as well as the optimal population for such testing.

In an effort to better characterize the potential benefit of pharmacogenomic testing in clinical populations with major depressive disorder (MDD), the present study aimed to assess the efficacy of assay-guided treatment (AGT) versus treatment-as-usual (TAU), as measured by change in SIGH-D-17 at 8 weeks, using a randomized, controlled, participant- and rater-blinded design across 21 U.S. sites. The investigators hypothesized that, consistent with similar studies, depression severity as measured by the SIGH-D-17 would be significantly reduced after 8 weeks.

2 | METHODS

This was an 8-week, multicenter, participant and rater-blinded randomized controlled trial. A screening visit was followed 1 week later by baseline visit (Week 0), with subsequent visits at Week 2, 4, 6, and 8. Eligible participants were of age 18–75 years, with a primary diagnosis of nonpsychotic MDD based on Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) criteria and MINI 7.0, and SIGH-D-17 score >18 (i.e., moderate to severe depression), at both screening and baseline visits. Participants were also required to have failure of at least one prior adequate trial of a standard antidepressant for the current major depressive episode (using Antidepressant Treatment Response Questionnaire [ATRQ] criteria—i.e., 6 weeks at adequate dose; Chandler, Iosifescu, Pollack, Targum, & Fava, 2010) due to inefficacy or intolerable adverse effects. Exclusion criteria included severe personality disorder traits (based on DSM-5 criteria) that, in the opinion of the site investigator, would interfere with the participation in the study or the evaluation of efficacy and safety, as well as all diagnosed personality disorders; current DSM-5 diagnosis of neurocognitive disorders, schizophrenia spectrum (lifetime diagnosis) and other psychotic disorders, bipolar and related disorders (lifetime diagnosis), trauma and stress-related disorders, obsessive compulsive disorder and related disorders. Participants with comorbid current anxiety disorders (except panic disorder) and adjustment disorders could be included if the site investigator considered MDD to be the primary diagnosis and the other disorders were assessed as stable. Conversely, the investigator could also exclude participants for any other DSM-5 disorders that, in the opinion of the site investigator, may have interfered with participation in the study or the evaluation of efficacy and safety.

Additional exclusion criteria included DSM-5 diagnosis of substance related and addictive disorders diagnosed in the last 12 months (other than tobacco and caffeine); history of suicidal behavior within 12 months of screening or presence of active suicidal ideation with intent in the past 12 months (Items 4 or 5) at screening or baseline, as determined by the Columbia Suicide Severity Rating Scale (C-SSRS), or individuals considered to be an acute suicide risk in the clinical judgment of the site investigator. Further severity exclusions included four or more failed pharmacologic interventions in the current major depressive episode (of which at least one must meet ATRQ criteria—i.e., 6 weeks at adequate dose); electroconvulsive therapy or transcranial magnetic stimulation therapy initiated within 90 days of screening or planned during the study period; or any psychotherapy, including cognitive behavioral therapy or dialectical behavioral therapy, initiated within 90 days of screening or planned during the study. Participants already receiving psychotherapy could continue treatment during the study period as long as frequency was not increased. Finally, participants were excluded if they reported unstable or active medical condition(s) which, in the opinion of the site investigator, would jeopardize the subject's safety or interfere with participation of the study or confound evaluation of efficacy or safety, including current diagnosis of uncontrolled hypothyroidism; females who were pregnant, nursing, or planning a pregnancy during the study were also excluded.

All participants signed written informed consent before study entry. The study protocol was approved by Western Institutional Review Board and posted to clinicaltrials.gov as NCT02634177. Of note, the master protocol was subsequently amended to cover enrollment of an additional 70 subjects ≥65 years of age in a second study using the same design, to be analyzed and reported separately.

3 | INTERVENTION

The Genecept Assay (version 2.0) is a commercially available test that incorporates 45 variants of 7 pharmacokinetic cytochrome P450 genes and 12 variants of 11 pharmacodynamic or other genes. Subjects were randomized 1:1 to either AGT or TAU treatment conditions. All subjects provided a DNA sample via buccal swab at the screening visit. In the AGT condition, assay results were provided via secure portal at baseline study visit to the treating investigator, who could have used the results to guide antidepressant pharmacotherapy; however, clinicians were permitted to prescribe drugs ad libitum, that is they were not constrained by the assay's results. In the TAU condition, the investigators treated the subjects without the knowledge of the pharmacogenetic testing results. Assay results for TAU subjects were provided to the investigator once all Week 8 visit

procedures had been completed. Subjects and raters utilizing the SIGH-D-17 were blinded to treatment condition; treating physicians were unblinded.

Before study initiation, training was provided to all participating investigators on the interpretation of genetic testing results and on the relevance of each genetic variant to pharmacotherapy. The training included the role of pharmacogenetics and the pharmacodynamic and pharmacokinetic relevance of each gene variation with respect to pharmacotherapy with psychotropic medications. In addition, during the study period, treating investigators could speak with pharmacists at Genomind regarding interpretation of pharmacogenetic testing results, but could not receive medication recommendations from the pharmacist.

4 | OUTCOMES

The primary efficacy outcome was change from Baseline in SIGH-D-17; secondary efficacy outcomes included Quick Inventory of Depressive Symptomatology-Self Report (QIDS-SR16) and Clinical Global Impression-Improvement (CGI-I) scores. Safety and tolerability outcomes included Frequency, Intensity, and Burden of Side Effects Rating (FIBSER), Patient-Rated Inventory of Side Effects, and Columbia-Suicide Severity Rating Scale (C-SSRS). All measures were collected at baseline/Week 0, 2, 4, 6, and 8; SIGH-D-17 and C-SSRS were also collected at screening. Structured blinded assessment of the SIGH-D-17 was conducted by site raters after undergoing a rater prequalification, training and certification process by an independent rater-training company.

5 | ANALYSIS

Analysis of change from baseline in SIGH-D-17 used mixed effects models with repeated measures (MMRM) in the full analysis set—that is all randomized subjects with a post-baseline SIGH-D-17 assessment. The model included the fixed effect continuous factor baseline SIGH-D-17, and fixed effect categorical factors investigative site, treatment group (AGT and TAU; 2 levels), visit (weeks 2, 4, 6 and 8; 4 levels), and treatment x visit interaction. AGT and TAU mean change in SIGH-D-17 at Week 8 was estimated and tested utilizing the -s (LS) means from the treatment x visit interaction in the MMRM model. The primary analysis tested the difference (contrast) between the Week 8 least squares (LS) means, at two-sided significance 0.05. Comparisons between the AGT and TAU means at Week 2, Week 4, and Week 6 were also generated for descriptive purposes.

Key continuous secondary endpoints were analyzed using the same MMRM approach as for the primary endpoint. Key categorical secondary endpoints (response, defined as 50% reduction from baseline SIGH-D; remission, defined as SIGH-D ≤ 7) were analyzed using the Mantel-Haenszel method, stratified on site.

Safety endpoints were analyzed in the full safety set—that is, all randomized subjects who complete the baseline appointment with the treating investigator—regardless of availability of a post-baseline visit.

Power estimate: The sample size calculation assumed a difference in SIGH-D-17 mean improvement scores = 3.1, a within-subject standard deviation (*SD*) = 7.2, a dropout rate at Week 8 of 23%, a two-sided 0.05 *t* test of treatment means, and 90% power. These assumptions result in a sample size of 150 randomized patients in each treatment group (300 total).

6 | RESULTS

Among 305 participants who completed the baseline visit, 304 were randomized, yielding 296 with evaluable outcomes (i.e., in the full analytic set; Figure 1).

Characteristics of the study population are summarized in Table 1. Among randomized participants, the most common medication additions included duloxetine (35, 11.5%), venlafaxine (35, 11.5%), bupropion (28, 9.2%); escitalopram (19, 6.3%); levomilnacipran (11, 3.6%), aripiprazole (10, 3.3%), desvenlafaxine (10, 3.3%), mirtazapine (9, 3.0%); sertraline (9, 3.0%), and vortioxetine (8, 2.6%). For the primary comparison of interest, change in SIGH-D-17, no significant difference was detected between AGT and TAU at Week 8 ($p = .53$; Table 2). Similarly, no significant differences in response (58/146 [39.7%] in AGT vs. 72/150 [48.0%] in TAU; −7.8% difference, $p = .17$) and remission (35/146 [24.0%] in AGT vs. 46/150 [30.7%] in TAU; −6.1% difference, $p = .23$) were identified (data not shown). Secondary efficacy measures including QIDS-SR (mean change −6.04 vs. −6.45; MMRM, 0.39) and CGI-I (mean change −1.74 vs. −1.65; MMRM, 0.56) yielded similar results (Tables S1 and S2). However, the proportion of CGI-I responders (score ≤3) was greater in the AGT group compared to TAU subjects (128/146, 87.7% vs. 118/150, 78.7%; $p = .036$).

Rates of study completion were also not different between the two groups (Figure 1). Among tolerability measures, no meaningful

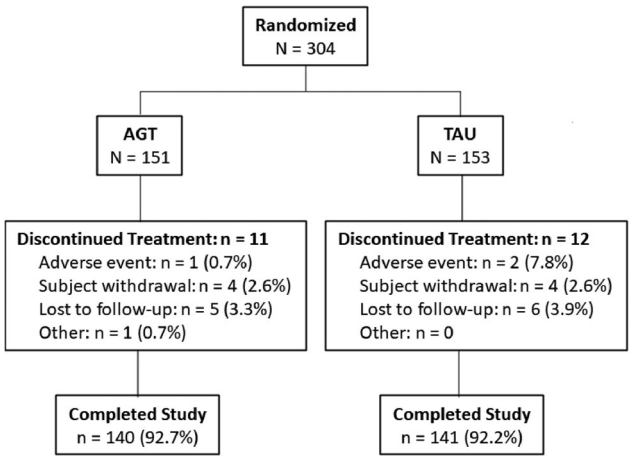

FIGURE 1 Subject disposition

TABLE 1 Baseline demographic and clinical features

Feature	Statistic/category	AGT (*N* = 151)	TAU (*N* = 153)	All Patients (*N* = 304)
Age (yrs)	Mean	47.8	47.6	47.7
	SD	12.38	12.06	12.20
Gender, *n* (%)	Male	44 (29.1)	42 (27.5)	86 (28.3)
	Female	107 (70.9)	111 (72.5)	218 (71.7)
Race, *n* (%)	Asian	0	1 (0.7)	1 (0.3)
	American Indian or Alaskan Native	2 (1.3)	1 (0.7)	3 (1.0)
	Black or African American	33 (21.9)	38 (24.8)	71 (23.4)
	Native Hawaiian or Other Pacific Islander	1 (0.7)	2 (1.3)	3 (1.0)
	White	111 (73.5)	110 (71.9)	221 (72.7)
	Other	4 (2.6)	1 (0.7)	5 (1.6)
Final SIGH-D-17 scores at Screening	Mean	22.1	22.3	22.2
	SD	3.23	3.21	3.22
Final SIGH-D-17 scores at baseline	Mean	22.5	22.1	22.3
	SD	3.41	3.24	3.32
Moderate vs. severe MDD, SIGH-D *n* (%)	<24, Moderate MDD[a]	90 (59.6)	106 (69.3)	196 (64.5)
	≥24, Severe MDD	61 (40.4)	47 (30.7)	108 (35.5)
QIDS-SR16 at baseline	Mean	14.8	14.6	14.7
	SD	4.15	3.86	4.00
CGI-S at baseline	Mean	4.3	4.3	4.3
	SD	0.50	0.49	0.49
Previous failed, adequate treatment >1[b]	0	1 (0.7)	0	1 (0.3)
	1	106 (70.2)	101 (66.0)	207 (68.1)
	2 or 3	44 (29.1)	51 (33.3)	95 (31.3)
	>3	0	1 (0.7)	1 (0.3)
	Antipsychotic	9 (6.0)	14 (9.2)	23 (7.6)
	Anxiolytic	2 (1.3)	2 (1.3)	4 (1.3)
	Misc. antidepressant	34 (22.5)	33 (21.6)	67 (22.0)
	Mood stabilizer	3 (2.0)	5 (3.3)	8 (2.6)
	SNRI	25 (16.6)	31 (20.3)	56 (18.4)
	SSRI	80 (53.0)	77 (50.3)	157 (51.6)
	TCA	3 (2.0)	5 (3.3)	8 (2.6)
BMI (kg/m^2) at screening	Mean	31.27	31.62	31.45
	SD	7.491	8.443	7.973

Note: Previous Depression Treatment is the MDD medication which started before baseline visit.

Abbreviations: AGT, assay-guided treatment; BMI, body mass index; CGI, Clinical Global Impression; MDD, major depressive disorder; QIDS-SR16, Quick Inventory of Depressive Symptomatology-Self Report; *SD*, standard deviation; SIGH-D-17, Hamilton Depression Rating Scale; SNRI, Serotonin–norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; TAU, treatment-as-usual; TCA, tricyclic antidepressant.

[a]The categories are summarized based on consensus scores at baseline. Denominator for percentages is the number of patients with non-missing baseline consensus SIGH-D-17 score.

[b]Trials of medications FDA approved for MDD or commonly used for MDD identified as failed medications, with daily dose greater than or equal to minimum labeled MDD dose for at least 6 weeks, as recorded on ATRQ.

differences were detected between AGT and TAU in terms of FIBSER score (Table S3).

In a post-hoc exploratory analysis, we sought to exclude the possibility that a subset of participants could fail to improve or even worsen with testing. At Week 8, significantly more individuals had failed to improve or worsened (by at least one point on the SIGH-D-17) in the TAU group; 17/186 (9.1%) versus 6/181 (3.3%) in the AGT group ($\chi^2 = 5.3$, $df = 1$, $p = .021$). Results were similar when worsening was defined as ≥1, ≥3, or ≥5 points greater than Baseline were included ($p = .007$, $p = .02$, $p = .037$, respectively).

Because investigators were not constrained by the assay's recommendations, an additional post-hoc analysis was performed in

which the prescribed treatment was defined as concordant or discordant with the assay's recommendations. Concordance was prospectively defined for each patient by a third-party reviewer without knowledge of outcomes, as nonconcordant, somewhat/mixed, clear concordance, or unable to categorize. Because genetic data were available for both treatment groups after the fact, we combined treatment groups and analyzed all subjects. Across both arms combined, patients obtaining treatment that was clearly concordant with pharmacogenetically informed evidence-based practice were more likely to remit compared to patients who received discordant or somewhat/mixed concordance combined (65/193 [33.7%] vs. 20/108 [18.5%]; odds ratio = 2.23, 95% CI: 1.17–2.83; [$\chi^2 = 7.85$, $df = 1$, $p = .005$]). Conversely, a post-hoc analysis identified no significant group-by-prior-treatment-trial-count interaction ($p = .361$), though such analyses must be interpreted with caution (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3068511/).

7 | DISCUSSION

In this randomized, controlled trial of 305 outpatient adults with nonpsychotic MDD, no significant difference in depression severity over 8 weeks was identified between the testing and the TAU arm. Secondary measures of depression, including QIDS-SR and Clinical Global Impression-Severity score, likewise did not distinguish intervention from control. There was no evidence that AGT led to poorer outcomes that TAU; in fact, an exploratory measure of worsening indicated significantly fewer participants worsened in the AGT group. An additional analysis suggests that individuals prescribed therapy concordant with the pharmacogenetic information provided may be more likely to remit.

These results, which contrast with two prior nonrandomized cost-effectiveness studies of this assay using health claims data (Fagerness et al., 2014; Perlis et al., 2018), must also be interpreted in the context of prior randomized trials, all single-blind and either 8 or 12 weeks. In particular, while two 12-week studies were positive (Singh, 2015), two recent studies yielded negative results on their primary outcome measures, including a recent study enrolling more than 1,100 participants (Greden et al., 2019; Pérez et al., 2017) These latter studies did demonstrate benefit on some secondary measures. Furthermore, a meta-analysis of five previously published trials of pharmacogenetic testing in individuals with depression found that such testing was associated with greater rates of remission (Bousman, Arandjelovic, Mancuso, Eyre, & Dunlop, 2019).

Faced with the high degree of similarity between these tests, the differing results are unlikely to be accounted for by substantive differences in assay design or reporting (Bousman & Dunlop, 2018). Notably, the present study investigated a similar study population to a recent large negative study, so it is possible that the results reflect a true absence of efficacy for these panels as currently configured (Greden et al., 2019). Even absent efficacy differences, the lack of apparent improvement in tolerability in this study and the recent GUIDED trial (Greden et al., 2019) suggests more work may be needed to understand the relationship between blood levels and tolerability. However, another potential contributor to the failure to detect efficacy is the substantial response rate of 48% among the TAU arm, larger than what would be expected from an MDD population with a history of one to three failures to respond to adequate antidepressant therapies. While the comparison is imperfect, a similar population in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial had an average response rate across levels 2 through 4 (i.e., 1–3 prior treatment failures) of 20.5% (Rush, Trivedi, & Wisniewski, 2006) consistent with more recent studies in treatment-resistant MDD (see, e.g., Papakostas [https://www.ncbi.nlm.nih.gov/pubmed/23212058] including a pharmacogenomic study, Greden et al., 2019). In essence, the high response rates in the TAU arm may have led to a ceiling effect, precluding detection of further benefit. In future studies, site-independent assessments and validation of the patients deemed appropriate for enrollment by the site investigators may be efficacious in addressing potential enrollment biases (Freeman, Pooley, & Flynn, 2017) Strategies to enable complete blinding (of participants, raters, and prescribers, rather than solely the first two as in this and prior studies) also merit consideration as a means of minimizing expectancy bias.

More generally, our results should prompt efforts to better understand differences in study design or study populations that may impact efficacy of pharmacogenetic tests. Ironically, one of the greatest challenges in precision medicine is identifying the population for which benefit is greatest. That is, where precisely does precision medicine work best? In particular, despite strong evidence that functional cytochrome P450 variation impacts blood levels, such variation remains uncommon—in Caucasian cohorts, ~10% of individuals (with the exception of CYP2C19) will exhibit non-wildtype metabolic phenotypes (Zanger & Schwab, 2013). Notably, prevalence of many of these variants is known to differ by ancestry, suggesting at least the possibility of differential efficacy among different populations. As such, in any given study, the benefit may accrue to this 10% of the population, while outcomes in the remaining 90% would resemble TAU. While preliminary and post-hoc, our finding of less likelihood of worsening in the AGT arm compared to TAU may suggest the utility of identifying high-risk individuals who may benefit most from testing. These individuals may be more likely to consume health care resources.

Another important next step will be understanding the extent to which individual aspects of pharmacogenomic tests may contribute to, or detract from, improvement in outcomes. Most such tests combine across pharmacokinetic and pharmacodynamic variants, with widely varying evidence bases underlying them. Despite the desire to aggregate such results into easily interpreted reports, such integration precludes efforts to identify the most and least effective elements of these tests. This understanding is critical in light of our post-hoc analyses suggesting that accurate report interpretation is a critical factor for successful application of pharmacogenetic analyses.

Finally, another key question is the optimal study duration; in particular, the present study does not address the potential for longer-term benefit associated with pharmacogenomic testing. To

TABLE 2 SIGH-D-17 total score values and percent change from baseline by visit (mixed-effects model for repeated measures)

Study visit	Statistic	AGT (*N* = 146)	TAU (*N* = 150)
Baseline	Mean (SE)	22.48 (0.284)	22.05 (0.264)
	SD	3.431	3.236
Week 2	Mean (SE)	17.39 (0.499)	17.77 (0.479)
	SD	5.948	5.766
Percent change from baseline to Week 2	Mean (SE)	−22.86 (1.986)	−19.69 (1.951)
	SD	23.668	23.496
	MMRM LSM (SE)[a]	−22.76 (1.925)	−19.60 (1.904)
	95% CI	(−26.55, −18.97)	(−23.34, −15.85)
	Diff in MMRM LSM (SE)[b]	−3.16 (2.719)	
	95% CI of Diff	(−8.51, 2.19)	
	MMRM *p* value[c]	0.2460	
Week 4	Mean (SE)	15.43 (0.570)	15.66 (0.531)
	SD	6.670	6.418
Percent change from baseline to Week 4	Mean (SE)	−31.44 (2.341)	−28.76 (2.299)
	SD	27.399	27.779
	MMRM LSM (SE)[a]	−31.74 (2.257)	−28.50 (2.205)
	95% CI	(−36.19, −27.30)	(−32.84, −24.16)
	Diff in MMRM LSM (SE)[b]	−3.24 (3.163)	
	95% CI of Diff	(−9.47, 2.98)	
	MMRM *p* value[c]	0.3060	
Week 6	Mean (SE)	13.93 (0.604)	14.02 (0.611)
	SD	7.038	7.172
Percent change from baseline to Week 6	Mean (SE)	−38.34 (2.572)	−36.12 (2.666)
	SD	29.992	31.317
	MMRM LSM (SE)[a]	−38.05 (2.509)	−35.34 (2.483)
	95% CI	(−42.99, −33.11)	(−40.23, −30.45)
	Diff in MMRM LSM (SE)[b]	−2.71 (3.538)	
	95% CI of Diff	(−9.68, 4.25)	
	MMRM *p* value[c]	0.4438	
Week 8	Mean (SE)	12.77 (0.566)	11.90 (0.565)
	SD	6.649	6.684
Percent change from baseline to Week 8	Mean (SE)	−43.34 (2.404)	−45.99 (2.537)
	SD	28.239	30.020
	MMRM LSM (SE)[a]	−43.29 (2.350)	−45.44 (2.329)
	95% CI	(−47.91, −38.66)	(−50.03, −40.86)
	Diff in MMRM LSM (SE)[b]	2.15 (3.316)	
	95% CI of Diff	(−4.37, 8.68)	
	MMRM *p* value[c]	0.5165	

Abbreviations: AGT, assay-guided treatment; CI, confidence interval; LSM, least squares mean; MMRM, mixed effects models with repeated measure; SD, standard deviation; SE, standard error; TAU, treatment-as-usual.

[a]In the MMRM model, the dependent variable is the change in SIGH-D-17 from baseline at each of the scheduled visits Week 2, Week 4, Week 6, and Week 8. See text for details.

[b]Difference between LSM changes for AGT and TAU (AGT−TAU) at the specified visit from MMRM analysis.

[c]Two-sided *p* value for treatment difference at specified visit from MMRM analysis.

date, all published studies report short-term results, with no blinded results beyond 12 weeks. Longer-term studies may be required to capture the impact of pharmacogenomic testing, particularly for diseases like MDD for which the costs of the disease accrue in poorer clinical outcomes more broadly. Such studies may also aid in reconciling results of randomized trials with those of longer-term cost-effectiveness investigations.

While our primary results do not support for the overall efficacy of pharmacogenetic testing in short-term antidepressant treatment, our post-hoc analyses combined with the observed secondary

endpoints of another trial provide hypotheses that merit further study: fewer patients worsen with assay guided therapy compared to treatment as usual, and provision of therapy that is actually concordant with the assay's pharmacogenetic guidance results in improved outcomes. Consistent with all prior studies, our results also support the safety of such testing; despite fears that testing might lead to use of less effective antidepressant strategies, none of the published literature find this to be the case. The paradox of economic and resource savings in the face of lack of overall efficacy in some prospective trials, requires further study, and may suggest that there are differences between real-world populations receiving pharmacogenetic testing and the population observed in prospective clinical trials. Efforts to better understand the subset of individuals who may derive benefit, and the time course over which such benefits may be identified, represent important next steps for psychiatric pharmacogenomic studies.

ACKNOWLEDGMENT

This study was funded by Genomind.

CONFLICT OF INTERESTS

Dr. Perlis has received consulting fees from Burrage Capital, Genomind, Outermost Therapeutics, RID Ventures, and Takeda. He holds equity in Outermost Therapeutics and Psy Therapeutics. Dr. Fava reports Research Support: Acadia Pharmaceuticals, Allergan, Alkermes, Inc., Aptinyx, Avanir Pharmaceuticals Inc., Axsome, Benckiser Pharmaceuticals, Inc., BioClinica, Inc., Biogen, BioHaven, Cambridge Science Corporation, Cerecor, Gate Neurosciences, Inc., GenOmind, LLC, Gentelon, LLC, Happify, Johnson & Johnson, Lundbeck Inc., Marinus Pharmaceuticals, Methylation Sciences, Inc., Millennium Pharmaceutics, Inc. Minerva Neurosciences, Neuralstem, NeuroRX Inc., Novartis, Otsuka, Pfizer, Premiere Research International, Relmada Therapeutics Inc., Reckitt, Shenox Pharmaceuticals, Stanley Medical Research Institute (SMRI), Taisho, Takeda, Vistagen, National Institute of Drug Abuse (NIDA); National Institutes of Health (NIH), National Institute of Mental Health (NIMH), and PCORI. Dr. Fava has not done any personal consulting. Any consulting he has done has been on behalf of Massachusetts General Hospital. Stock/Other Financial Options: Equity Holdings: Compellis; Psy Therapeutics. Royalty/patent, other income: Patents for Sequential Parallel Comparison Design (SPCD), licensed by MGH to Pharmaceutical Product Development, LLC (PPD) (US_7840419, US_7647235, US_7983936, US_8145504, US_8145505); and patent application for a combination of ketamine plus scopolamine in major depressive disorder (MDD), licensed by MGH to Biohaven. Patents for pharmacogenomics of Depression Treatment with Folate (US_9546401, US_9540691). Copyright for the MGH Cognitive & Physical Functioning Questionnaire (CPFQ), Sexual Functioning Inventory (SFI), Antidepressant Treatment Response Questionnaire (ATRQ), Discontinuation-Emergent Signs and Symptoms (DESS), Symptoms of Depression Questionnaire (SDQ), and SAFER; Lippincott, Williams and Wilkins; Wolkers Kluwer; World Scientific Publishing Co. Pte. Ltd. Drs. Krause and Dowd are employees of Genomind, the study's sponsor. Dr. Lencz received consulting fees from Genomind for aid in data analysis of the study.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available to collaborating investigators from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Roy H. Perlis http://orcid.org/0000-0002-5862-6757
David S. Krause http://orcid.org/0000-0002-2008-1974

REFERENCES

Bousman, C. A., Arandjelovic, K., Mancuso, S. G., Eyre, H. A., & Dunlop, B. W. (2019). Pharmacogenetic tests and depressive symptom remission: A meta-analysis of randomized controlled trials. *Pharmacogenomics*, *20*(1), 37–47.

Bousman, C. A., & Dunlop, B. W. (2018). Genotype, phenotype, and medication recommendation agreement among commercial pharmacogenetic-based decision support tools. *Pharmacogenomics Journal*, *18*(5), 613–622. https://doi.org/10.1038/s41397-018-0027-3

Bradley, P., Shiekh, M., Mehra, V., Vrbicky, K., Layle, S., Olson, M. C., … Lukowiak, A. A. (2018). Improved efficacy with targeted pharmacogenetic-guided treatment of patients with depression and anxiety: A randomized clinical trial demonstrating clinical utility. *Journal of Psychiatric Research*, *96*, 100–107. https://doi.org/10.1016/j.jpsychires.2017.09.024

Cacabelos, R., Cacabelos, N., & Carril, J. C. (2019). The role of pharmacogenomics in adverse drug reactions. *Expert Review of Clinical Pharmacology*, *12*(5), 407–442.

Chandler, G. M., Iosifescu, D. V., Pollack, M. H., Targum, S. D., & Fava, M. (2010). RESEARCH: Validation of the Massachusetts General Hospital Antidepressant Treatment History Questionnaire (ATRQ): Validation of the MGH ATRQ. *CNS Neuroscience & Therapeutics*, *16*(5), 322–325. https://doi.org/10.1111/j.1755-5949.2009.00102.x

Fagerness, J., Fonseca, E., Hess, G. P., Scott, R., Gardner, K. R., Koffler, M., … Lombard, J. (2014). Pharmacogenetic-guided psychiatric intervention associated with increased adherence and cost savings. *The American Journal of Managed Care*, *20*(5), e146–e156.

Food and Drug Administration. (2018). Table of pharmacogenomic biomarkers in drug labeling. Retreived from https://www.fda.gov/downloads/Drugs/ScienceResearch/UCM578588.pdf

Freeman, M. P., Pooley, J., Flynn, M. J., Baer, L., Mischoulon, D., Mou, D., & Fava, M. (2017). Guarding the gate: Remote structured assessments to enhance enrollment precision in depression trials. *Journal of Clinical Psychopharmacology*, *37*(2), 176–181.

Greden, J. F., Parikh, S. V., Rothschild, A. J., Thase, M. E., Dunlop, B. W., DeBattista, C., … Dechairo, B. (2019). Impact of pharmacogenomics on clinical outcomes in major depressive disorder in the GUIDED trial: A large, patient- and rater-blinded, randomized, controlled study. *Journal of Psychiatric Research*, *111*, 59–67. https://doi.org/10.1016/j.jpsychires.2019.01.003

McCoy, T. H., Castro, V. M., Cagan, A., Roberson, A. M., & Perlis, R. H. (2017). Prevalence and implications of cytochrome P450 substrates in Massachusetts hospital discharges. *Pharmacogenomics Journal*, *17*(4), 382–385. https://doi.org/10.1038/tpj.2016.24

Pérez, V., Salavert, A., Espadaler, J., Tuson, M., Saiz-Ruiz, J., Sáez-Navarro, C., … Menchón, J. M. (2017). Efficacy of prospective pharmacogenetic testing in the treatment of major depressive disorder: Results of a randomized, double-blind clinical trial. *BMC Psychiatry*, *17*(1), 250. https://doi.org/10.1186/s12888-017-1412-1

Perlis, R. H., Mehta, R., Edwards, A. M., Tiwari, A., & Imbens, G. W. (2018). Pharmacogenetic testing among patients with mood and anxiety disorders is associated with decreased utilization and cost: A propensity-score matched study. *Depression and Anxiety*, *35*(10), 946–952. https://doi.org/10.1002/da.22742

Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., … Fava, M. (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *American Journal of Psychiatry*, *163*(11), 1905–1917.

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

Zanger, U. M., & Schwab, M. (2013). Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacology and Therapeutics*, *138*(1), 103–141.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Perlis RH, Dowd D, Fava M, Lencz T, Krause DS. 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. https://doi.org/10.1002/da.23029