Impact of pharmacogenomics on clinical outcomes in major depressive disorder in the GUIDED trial: A large, patient- and rater-blinded, randomized, controlled study

John F. Greden[a,*], Sagar V. Parikh[a], Anthony J. Rothschild[b], Michael E. Thase[c], Boadie W. Dunlop[d], Charles DeBattista[e], Charles R. Conway[f], Brent P. Forester[g], Francis M. Mondimore[h], Richard C. Shelton[i], Matthew Macaluso[j], James Li[k], Krystal Brown[l], Alexa Gilbert[k], Lindsey Burns[k], Michael R. Jablonski[k], Bryan Dechairo[k,l]

[a] University of Michigan Department of Psychiatry and Comprehensive Depression Center 4250 Plymouth Rd, Ann Arbor, MI, 48109, USA
[b] University of Massachusetts Medical School and UMass Memorial Healthcare, 55 N Lake Ave, Worcester, MA, 01655, USA
[c] Perelman School of Medicine of the University of Pennsylvania and the Corporal Michael Crescenz VAMC, 3400 Civic Center Blvd, Philadelphia, PA, 19104, USA
[d] Emory University School of Medicine, Department of Psychiatry and Behavioral Sciences, 12 Executive Park Dr NE #200, Atlanta, GA, 30329, USA
[e] Stanford University School of Medicine, Department of Psychiatry and Behavioral Sciences, 401 Quarry Rd, Stanford, CA, 94305, USA
[f] Washington University School of Medicine, Department of Psychiatry, The John Cochran Veteran's Administration Hospital, 660 S Euclid Ave, St. Louis, MO, 63110, USA
[g] McLean Hospital, Division of Geriatric Psychiatry, Harvard Medical School, 115 Mill St, Belmont, MA, 02478, USA
[h] Johns Hopkins University School of Medicine, Department of Psychiatry and Behavioral Sciences, 1800 Orleans St, Baltimore, MD, 21287, USA
[i] The University of Alabama at Birmingham, Department of Psychiatry and School of Medicine, 1720 2nd Ave S, Birmingham, AL, USA
[j] University of Kansas School of Medicine-Wichita, Department of Psychiatry and Behavioral Sciences, 1010 N Kansas St, Wichita, KS, 67214, USA
[k] Assurex Health, Inc., 6960 Cintas Blvd, Mason, OH, 45040, USA
[l] Myriad Genetics, Inc., 320 Wakara Way, Salt Lake City, UT, 84108, USA

A B S T R A C T

Current prescribing practices for major depressive disorder (MDD) produce limited treatment success. Although pharmacogenomics may improve outcomes by identifying genetically inappropriate medications, studies to date were limited in scope. Outpatients (N = 1167) diagnosed with MDD and with a patient- or clinician-reported inadequate response to at least one antidepressant were enrolled in the **G**enomics **U**sed to **I**mprove **DE**pression **D**ecisions (GUIDED) trial – a rater- and patient-blind randomized controlled trial. Patients were randomized to treatment as usual (TAU) or a pharmacogenomics-guided intervention arm in which clinicians had access to a pharmacogenomic test report to inform medication selections (guided-care). Medications were considered congruent ('use as directed' or 'use with caution' test categories) or incongruent ('use with increased caution and with more frequent monitoring' test category) with test results. Unblinding occurred after week 8. Primary outcome was symptom improvement [change in 17-item Hamilton Depression Rating Scale (HAM-D17)] at week 8; secondary outcomes were response (≥50% decrease in HAM-D17) and remission (HAM-D17 ≤ 7) at week 8. At week 8, symptom improvement for guided-care was not significantly different than TAU (27.2% versus 24.4%, p = 0.107); however, improvements in response (26.0% versus 19.9%, p = 0.013) and remission (15.3% versus 10.1%, p = 0.007) were statistically significant. Patients taking incongruent medications prior to baseline who switched to congruent medications by week 8 experienced greater symptom improvement (33.5% versus 21.1%, p = 0.002), response (28.5% versus 16.7%, p = 0.036), and remission (21.5% versus 8.5%, p = 0.007) compared to those remaining incongruent. Pharmacogenomic testing did not significantly improve mean symptoms but did significantly improve response and remission rates for difficult-to-treat depression patients over standard of care (ClinicalTrials.gov NCT02109939).

1. Introduction

Major depressive disorder (MDD) is a serious and prevalent illness that continues to be associated with world-leading morbidity and disability as determined by the World Health Organization. (Organization, 2017). Current treatment approaches rely primarily on antidepressant medications and psychotherapy based on clinician's choice, preference, and experience. Studies consistently document that approximately half of patients with moderate-to-severe MDD do not respond adequately to their first medication, as originally shown in the pivotal STAR*D trial where only 49% of first-episode patients achieved response and 37% achieved remission after the first line of treatment (Rush et al., 2006b).

If patients failed to respond to this first intervention, the struggle to achieve response and remission continued during subsequent medication trials, with only 13% of patients achieving remission during their fourth medication (Rush et al., 2006b). Ultimately, current treatment approaches result in a long duration of unremitted illness for many patients, which is associated with worsened long-term prognosis (Ghio et al., 2014), adverse changes in brain function and morphology (Ghio et al., 2014), increased side-effect burden (Mrazek et al., 2014), intensification of other medical illnesses and corresponding annual medical costs (Mrazek et al., 2014), increased suicide risk, and adverse occupational and financial consequences (Greden, 2013; Trivedi et al., 2013). These data highlight the need for more effective treatment approaches.

Genetic variation is one of a number of variables that may impact the efficacy of a medication by affecting its metabolism (pharmacokinetics) or mechanism of action (pharmacodynamics). Medication safety may also be impacted by gene-drug interactions that result in harmful side-effects (Zhou et al., 2015). Pharmacogenomic assessments emerged as data-driven strategies to develop a more personalized, precise approach to antidepressant selection. Available strategies initially consisted of testing for single genes that encode for selected cytochrome P450 (CYP450) enzymes, which metabolize most psychotropic medications. As the field evolved, additional genes related to pharmacokinetics and pharmacodynamics were investigated. This multi-gene testing approach continued to utilize individual genotypes to determine single gene-drug interactions for each evaluated gene (Bradley et al., 2018; Perez et al., 2017; Singh, 2015). A recent randomized controlled trial found that multi-gene testing improved response and remission rates (Bradley et al., 2018) and reduced the economic burden among patients with very severe depression (n = 93) (Maciel et al., 2018).

While multi-gene tests provide added single genotype information, they do not assess the combined, interactive effects of multiple genes and genotypes (Swen et al., 2018). For instance, drugs may be metabolized through multiple pathways such that reduced function in one gene can be offset by increased function in another. To illustrate, for citalopram, decreased CYP2D6 function in an individual might be balanced by increased function of CYP2C19 to result in normal metabolism (Hicks et al., 2015). Depending on the combined phenotype with CYP2C19, a CYP2D6 poor metabolizer may have normal, lower or higher exposure to citalopram than population means. To address such considerations, a pharmacogenomic test was developed to genotype multiple relevant genes and produce a combined phenotype. The aim was to weight the genotypes for all measured variants or alleles in order to rank each medication according to the combined effect of all gene-drug interactions that impact the efficacy and safety of the drug for an individual patient (Hall-Flavin et al., 2012). The apparent clinical utility (Hall-Flavin et al., 2013; Hall-Flavin et al., 2012; Winner, J.G. et al., 2013), analytical validity (Jablonski et al., 2018), and cost effectiveness (Brown et al., 2017; Hornberger et al., 2015; (Jablonski et al., 2018; Winner, J. et al., 2013; Winner et al., 2015) of this weighted and combined, multi-gene pharmacogenomic test have been shown in patients with depression who have failed one or more medication.

The number of commercially available pharmacogenomic tests has grown; however, there are also growing uncertainties and concerns regarding the clinical utility of such tests (Goldberg, 2017; Rosenblat et al., 2017; Zeier et al., 2018; Zubenko et al., 2018). The most commonly expressed concerns and criticisms have pertained to study design, as evidence supporting the efficacy of pharmacogenomic testing has stemmed largely from non-randomized open-label trials or small randomized trials (Goldberg, 2017; Rosenblat et al., 2017; Zeier et al., 2018; Zubenko et al., 2018). There also have been concerns regarding study duration, cohort composition, and the incorporation of a robust comparator arm (Goldberg, 2017; Rosenblat et al., 2017; Zeier et al., 2018; Zubenko et al., 2018). Additionally, a recent study of pharmacogenomic tests found that there is a high level of disagreement between tests, suggesting that pharmacogenomic tests are not interchangeable and likely require separate evaluations to determine clinical utility (Bousman and Dunlop, 2018).

To concomitantly address many of the expressed study design concerns while evaluating the utility of pharmacogenomic testing in patients with MDD, we present data from the **G**enomics **U**sed to **I**mprove **DE**pression **D**ecisions (GUIDED) trial. To the authors' knowledge, this is the largest, blinded, randomized, long-term, controlled trial using pharmacogenomic testing for MDD to date. In this trial, active treatment guided by pharmacogenomic testing was compared to unguided active treatment (treatment as usual, TAU) in patients with MDD who had failed to respond to at least one adequate prior medication trial. Symptom improvement, response, and remission were monitored over 24 weeks with the primary endpoint at week 8.

2. Material and methods

2.1. Pharmacogenomic testing

The GeneSight® Psychotropic test from Assurex Health, Inc. (Mason, OH) was used for pharmacogenomic testing on all patients, as previously described (Jablonski et.al., 2018). The genotypes of 59 alleles and variants across 8 genes (CYP1A2, CYP2C9, CYP2C19, CYP3A4, CYP2B6, CYP2D6, HTR2A, SLC6A4; Supplemental Table 1) were evaluated by a pharmacogenomic algorithm. This is a proprietary test based on licensed technology disclosed in issued patents (U.S. patent no. 8,401,801 and U.S. patent no. 8,688,385). In brief, the algorithm weighed the combined influence of each individual genotype on patient response to each individual medication (Hall-Flavin et al., 2012). Based on this weighted and combined phenotype, 38 psychotropic medications were categorized based on three levels of gene-drug interaction: 'use as directed' (no detected gene-drug interactions), 'use with caution' (moderate gene-drug interactions; i.e. medications may be effective with dose modification), 'use with increased caution and with more frequent monitoring' (severe gene-drug interactions that may significantly impact drug safety and/or efficacy).

2.2. Study description

The GUIDED trial was a 24-week, randomized, controlled trial that evaluated outcomes when a pharmacogenomic test was used to guide medication selection (guided-care) compared to TAU. The trial protocol was approved by the Copernicus Group independent review board (INC1-14-012) and was performed in accordance with the principles of the Declaration of Helsinki. All patients provided written informed consent after receiving a complete description of the study.

The full study design is described in the Supplemental Methods. In brief, patients were enrolled at screening between April 14, 2014 and February 10, 2017. Eligible patients were randomized 1:1 to TAU or the guided-care (intervention) arm. Pharmacogenomic testing was performed for all patients between screening and baseline visits.

This trial differed from traditional drug studies in that patients in both arms received active treatment. For patients in TAU, active treatment was unguided by pharmacogenomic data. For patients in the guided-care arm, clinicians had access to the pharmacogenomic test report prior to the baseline visit to use in medication selection. Patients and raters were blinded to study arm. Clinicians could not be blinded to study arm, as the pharmacogenomic test results had to be consulted to select potential medications for the guided-care arm. Adherence to the test results was not mandated. For patients in TAU, clinicians were blinded to the pharmacogenomic test result until *after* completion of the week 8 visit. Sites were instructed to unblind patients in *both* arms after assessments at week 12 were completed. Because unblinding may have occurred prior to week 12 assessments, data collected through week 8 were considered blinded.

Patient assessments were performed at weeks 0 (baseline), 4, 8, 12, and 24. The primary assessment was the 17-item Hamilton Depression

rating scale (HAM-D17), administered by a blinded central rater (MedAvante-ProPhase Inc., Hamilton, NJ). Secondary assessments included the 16-item Quick Inventory of Depression Symptomology (QIDS-C16; administered by a blinded site rater) and 9-item Patient Health Questionnaire (PHQ-9; patient reported). No assessments completed by the unblinded clinician are included here.

2.3. Participants

Outpatients were enrolled from 60 academic and community sites in the U.S. that included psychiatric and primary care providers. Patients were included if they were over the age of 18, were diagnosed with MDD (≥ 11 on the QIDS-C16 and self-rated QIDS-SR16 at screening and baseline) and had an inadequate response (lack of clinical improvement or intolerable side-effects as reported by the patient or treating clinician) to at least one documented psychotropic treatment included on the pharmacogenomic test report within the current depressive episode. Patients were excluded if they had significant suicidal risk, some severe co-occurring psychiatric or cognitive disorders, and/or unstable or significant medical conditions, as described in the Supplemental Methods.

Analyses were conducted for two pre-specified cohorts. The intent-to-treat cohort included patients who met eligibility criteria. The per-protocol cohort additionally excluded patients who had a score of < 14 on the HAM-D17 at baseline to ensure the exclusion of patients with mild depression. Patients also were excluded if there were protocol violations or if the clinician did not view the electronic test report prior to the baseline visit (guided-care arm only).

2.4. Medication congruence with pharmacogenomic testing

Consistent with previous studies, prescribed medications were considered *congruent* with the pharmacogenomic test if they were in the 'use as directed' or 'use with caution' report categories (Altar et al., 2015; Winner et al., 2015). Although medications in the 'use with caution' category were subject to moderate gene-drug interactions, they were considered congruent because minor clinical modifications (e.g. dose adjustment) were predicted to make medications in this category safe and/or effective. Medications were considered *incongruent* if they were in the 'use with increased caution and with more frequent monitoring' category.

A post-hoc analysis evaluated the sub-set of patients in both study arms taking ≥ 1 incongruent medication prior to medication changes at baseline. Outcomes were evaluated for patients who changed to congruent medications by week 8 versus those who remained on incongruent medications. Patients were considered as taking congruent medications if none of their prescribed medications were incongruent. For TAU, this included patients who were incidentally switched to congruent medications without the aid of the pharmacogenomic test report.

2.5. Side effects

Patient-reported side effects were assessed as the mean number of side effects and proportion of patients reporting side effects according to study arm and baseline medication congruence. Only side effects with a probability of causal link to medication (e.g. categorized as likely, probably, possibly, or definitely relating to medication) were included.

2.6. Statistical analysis

Analyses were performed according to a pre-specified statistical analysis plan. The primary outcome was *symptom improvement* at week 8 in the per-protocol cohort, as measured by the percent change in HAM-D17 (blinded central rater) from baseline. Secondary outcomes

included response and remission at week 8 according to HAM-D17. Patient outcomes were further assessed by QIDS-C16 (blinded site rater), and PHQ-9 (blinded patient rated). *Response* was defined as $\geq 50\%$ decrease at week 8 in the assessment of interest (HAM-D17, QIDS-C16 or PHQ-9) from baseline. *Remission* was defined as having a score of ≤ 7 for HAM-D17, ≤ 5 for QIDS-C16, and < 5 for PHQ-9.

Patient demographics and clinical characteristics were assessed using descriptive statistics. The severity of depression was categorized according to HAM-D17 scores: 0-7, normal; 8-13, mild depression; 14-18, moderate depression; 19-22, severe depression; ≥ 23, very severe depression. A post-hoc moderator analysis was performed to determine whether the following baseline variables significantly affected the primary outcome: age, gender, HAM-D17 score, genetic congruence of medications, pharmacogenomic report category, number of failed medications, race (Caucasian, non-Caucasian), and ethnicity (Hispanic, non-Hispanic).

Analyses were performed for patients who completed the study through week 8. A Mixed Model for Repeated Measures (MMRM) was used for percentage change from baseline in HAM-D17, QIDS-C16, and PHQ-9 analyses, and generalized linear mixed model was used for response and remission analyses. T-tests were used to obtain two-sided p-values for study arms and comparison groups, where values ≤ 0.05 were considered statistically significant. Analyses were performed with SAS software (version 9.4) or JMP 13 (SAS Institute).

3. Results

3.1. Cohort description

Overall, 2004 patients were screened for inclusion in this study; 1799 met eligibility criteria and were randomized to TAU or guided-care after pharmacogenomic testing (Supplemental Fig. 1). After randomization, an additional 258 patients were excluded because they didn't meet the eligibility criteria at baseline or were lost to follow-up prior to baseline. A total of 1541 patients completed the baseline visit and were included in the final intent-to-treat cohort. For the per-protocol cohort, an additional 143 patients were excluded; 101 patients had a baseline HAM-D17 score < 14 (below mild depression) and 42 patients had protocol violations. The final per-protocol cohort included 1398 patients (717 patients in TAU, 681 patients in guided-care).

Although a minimum of one failed medication trial was required for enrollment, the mean number of failed medications was 3.51 (Table 1). The mean HAM-D17 score at baseline was 21.28, with a relatively equal distribution of patients with moderate (392/1398, 28.0%), severe (493/1398, 35.3%), or very severe (513/1398, 36.7%) depression (Table 1). General anxiety disorder was the most common psychiatric comorbidity. The mean age was 47.5 years and the majority (987/1398, 70.6%) of patients were female (Table 1). There were no substantial differences in demographics or disease by treatment arm (Table 1) or between the per-protocol and intent-to-treat cohorts (Supplemental Table 2).

At baseline, only 18.3% (256/1398) of patients were taking medications in the most severe report category ('use with increased caution and with more frequent monitoring'). The majority of patients were taking medications in the 'use as directed' (357/1398, 25.5%) or 'use with caution' (575/1398, 41.1%) report categories (Table 1). The remaining 15.0% (210/1398) of patients were not taking any medications included on the pharmacogenomic test report at baseline. There were no substantial differences between the individual gene phenotypes as determined by the pharmacogenomic test by study arm (Supplemental Table 3). Selective serotonin reuptake inhibitors (SSRIs) were the most commonly prescribed psychotropic medications at baseline for both the guided-care arm and TAU, accounting for nearly half of all prescribed medications (Supplemental Table 4).

In the per-protocol cohort, 1167 patients completed the study through the blinded week 8 endpoint (607 patients in TAU, 560

Table 1
Baseline demographics for patients in the per-protocol cohort who completed the baseline visit.

Characteristic	Treatment Arm				Total (N = 1398)	
	TAU (N = 717)		Guided–Care (N = 681)			
	N	%	N	%	N	%
Age Group						
18–34 years	158	22.0	162	23.8	320	22.9
35–49 years	192	26.8	200	29.4	392	28.0
50–64 years	266	37.1	235	34.5	501	35.8
65 years and over	101	14.1	84	12.3	185	13.2
Sex						
Female	498	69.5	489	71.8	987	70.6
Male	219	30.5	192	28.2	411	29.4
Ethnicity						
Hispanic or Latino	54	7.5	57	8.4	111	7.9
Not Hispanic or Latino	663	92.5	624	91.6	1287	92.1
Race						
White	589	82.1	538	79.0	1127	80.6
Black	94	13.1	114	16.7	208	14.9
Asian	17	2.4	12	1.8	29	2.1
American Indian or Alaska Native	3	0.4	5	0.7	8	0.6
Native Hawaiian or Other Pacific Islander	0	0	1	0.1	1	0.1
Other or Multiple	14	2.0	11	1.6	25	1.8
Depression Category						
Moderate (HAM-D17 14-18)	187	26.1	205	30.1	392	28.0
Severe (HAM-D17 19-22)	264	36.8	229	33.6	493	35.3
Very Severe (HAM-D17 ≥ 23)	266	37.1	247	36.3	513	36.7
Psychiatric Comorbidities						
General anxiety disorder	96	13.4	116	17.0	212	15.2
Panic disorders/social phobia	108	15.1	104	15.3	212	15.2
Post-traumatic stress disorder	32	4.5	36	5.3	68	4.9
Pharmacogenomic Report Category[a]						
Use as Directed	181	25.2	176	25.8	357	25.5
Use with Caution	295	41.1	280	41.1	575	41.1
Use with Increased Caution and with More Frequent Monitoring	138	19.2	118	17.3	256	18.3
Not Applicable[b]	103	14.4	107	15.7	210	15.0
	Mean (SD)	**Min, Max**	**Mean (SD)**	**Min, Max**	**Mean (SD)**	**Min, Max**
Age (years)	48.0 (14.5)	18, 85	46.9 (14.5)	18, 90	47.5 (14.5)	18, 90
HAM-D17 Score	21.4 (4.22)	14, 35	21.1 (4.20)	14, 37	21.3 (4.21)	14, 37
Failed Medication Trials	3.53 (3.01)	1, 34	3.48 (3.09)	1, 25	3.51 (3.05)	0, 34

[a] For patients taking more than one medication at baseline, the most severe report category is included.
[b] Includes patients not taking any medications included on the pharmacogenomic test report at baseline.

patients in the guided-care arm). A total of 913 patients completed the study through week 24, with weeks 12–24 being unblinded (456 patients in TAU, 457 patients in the guided-care arm; Supplemental Fig. 1). There were no substantial differences in the baseline demographics according to study arm for those who completed week 8 (Supplemental Table 5). Use of SSRIs decreased from baseline to week 8 while use of other psychotropic medications increased (Supplemental Table 4).

3.2. Symptom improvement, response, and remission

At week 8, there was a 27.2% decrease in HAM-D17 scores in the guided-care arm compared to a 24.4% decrease in TAU (Fig. 1). This difference in the primary outcome was not significant (p = 0.107). Differences in the key secondary outcomes of response and remission were positive and significant. The response rate among patients in the guided-care arm was 26.0% (146/560) at week 8, which was significantly higher than in TAU [19.9% (121/607), p = 0.013; Fig. 1]. Similarly, the rate of remission among patients in the guided-care arm (15.3%, 86/560) was significantly higher than in TAU [10.1% (61/607), p = 0.007; Fig. 1]. Although overall improvement was modest, this represents a 30% improvement in response, and 50% improvement in remission among these difficult-to-treat patients in the guided-care arm compared to TAU.

A moderator analysis was performed to assess whether any clinical or demographic variables affected treatment performance for the primary outcome of symptom improvement. There were no significant moderation factors (Table 2), supporting that the primary outcome according to study arm was not dependent on the other factors evaluated.

Continuous changes in HAM-D17 score from baseline to week 8 were assessed to evaluate why the continuous endpoint of symptom improvement did not reach statistical significance while the categorical endpoints (response, remission) were significant. This revealed that the distribution of continuous HAM-D17 score improvement from baseline to week 8 was shifted towards extreme improvement (≥50% decrease in HAM-D17; definition of response) in the guided care arm and towards modest improvement in TAU (Supplemental Fig. 2). As a result, the mean HAM-D17 improvement was similar for both study arms (Δ2.8%, p = 0.107) while the proportion of patients with extreme improvement in the guided-care arm drove a significant difference in the rate of response and remission.

Symptom improvement, response, and remission based on HAM-D17 in the intent-to-treat cohort (Fig. 2) were similar to the per-protocol cohort (Fig. 1). Therefore, outcomes based on secondary rating scales (QIDS-C16, blinded-site rater; PHQ-9, blinded-patient rated) were assessed in the intent-to-treat cohort. At week 8, symptom improvement (QIDS-C16 p = 0.182; PHQ-9 p = 0.036), response (QIDS-
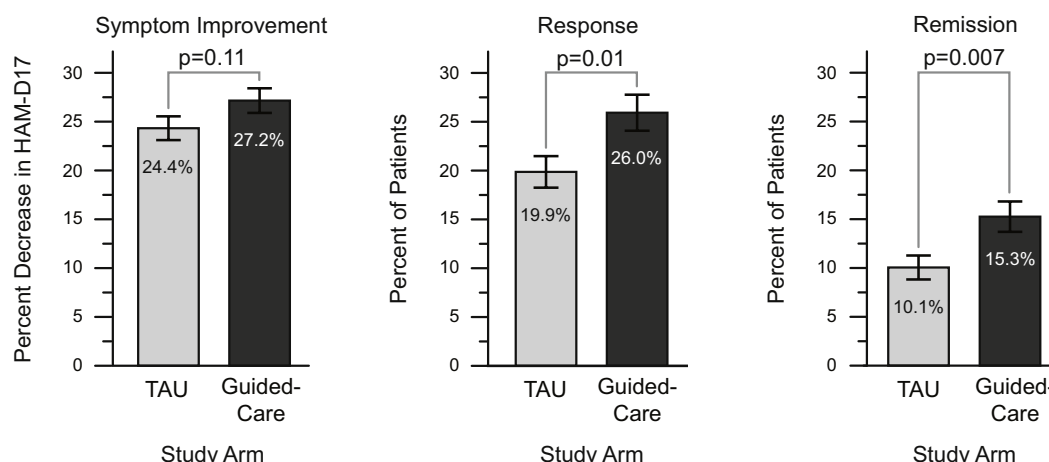
Fig. 1. Patient outcomes at week 8 in the pharmacogenomics guided-care arm (n = 560) compared to treatment as usual (n = 607). Outcomes were evaluated using the HAM-D17 depression rating scales.

Table 2
Moderator analysis evaluating the impact of clinical variables at baseline on the primary outcome (symptom improvement by treatment arm).

Variable	t-value	p-Value[a]
Age[b]	−0.86	0.3877
Gender	−0.31	0.7531
HAM-D17 Score at Baseline[b]	1.00	0.3168
Congruency of Baseline Medications (Congruent, Incongruent)	1.08	0.2802
PGx Report Category of Baseline Medications[d]	0.65[c]	0.5204
Number of Failed Medications at Baseline[b]	0.99	0.3221
Race (Caucasian, non-Caucasian)	−0.75	0.4538
Ethnicity (Hispanic, non-Hispanic)	−1.63	0.1033

[a] The percentage change in HAM-D17 at week 8 from baseline was analyzed using separate models that included treatment arm, baseline HAM-D17 score, the variable of interest (e.g., age), and treatment arm-by-variable of interest. Reported p-values are from the tests of moderation effect - treatment arm-by-variable of interest. P-values less than 0.05 indicate a significant effect of the clinical variable on treatment arms.

[b] Evaluated as a continuous variable.

[c] F-test statistic.

[d] Report categories were 'use as directed', 'use with caution', or 'use with increased caution and with more frequent monitoring'.

C16 p = 0.285; PHQ-9 p = 0.002), and remission (QIDS-C16 p = 0.014; PHQ-9 p = 0.066) were improved in the guided-care arm relative to TAU, though not all differences were significant (Fig. 2).

Patient outcomes were evaluated over the full 24-week study period in the guided-care arm of the per-protocol cohort to assess the longer-term impact of pharmacogenomic testing. Sustained decreases on the HAM-D17 rating scale demonstrated durable and continuing symptom improvement after unblinding, with HAM-D17 scores decreasing by 42.5% at week 24 relative to baseline (Fig. 3). This represents a 50% improvement from week 8 outcomes. The rates of response and remission increased by 70% and 100%, respectively, from week 8 to week 24. Specifically, 44.3% (203/457) of patients experienced response and 31.1% (142/457) of patients experienced remission at week 24 (Fig. 3).

3.3. Congruence with pharmacogenomic testing and outcomes

Prior to treatment changes at baseline, 79.4% (456/574) of patients in the guided-care arm and 77.5% (476/614) of patients in TAU were prescribed medications that were congruent with the pharmacogenomic test report (Supplemental Fig. 3). The proportion of patients who were prescribed congruent medications at week 8 increased to 91.2% (508/557) for the guided-care arm and remained relatively unchanged

in TAU (Supplemental Fig. 3). As already noted, the prescription of genetically congruent medications in TAU was incidental and not based on the pharmacogenomic test report.

One of the greatest potential utilities of the pharmacogenomic test is for patients taking incongruent medications, i.e., those with a potentially deleterious gene-drug interaction. An analysis of patients in both arms who entered the study on incongruent medications was performed. At baseline, patients on incongruent (N = 213) and congruent medications (N = 805) had similar mean HAM-D17 scores (21.19 and 21.18, respectively) and number of failed medication trails (3.77 and 3.59, respectively) (Supplemental Table 6). The analysis of patients on incongruent medications at baseline showed that outcomes were significantly improved among those who switched to a congruent medication by week 8. There was a 33.5% decrease in HAM-D17 scores among patients who switched to congruent medications compared to 21.1% among those who remained on incongruent medications (p = 0.002; Fig. 4). In addition, this analysis showed statistically significant improvements among patients on congruent versus incongruent medications at week 8 in the rates of response [28.5% (22/77) versus 16.7% (23/136), p = 0.036] and remission [21.5% (17/77) versus 8.5% (12/136), p = 0.007].

3.4. Side effects

There were no statistically significant differences between the guided-care arm and TAU regarding the mean number of side effects at week 8 (0.243 versus 0.237, p = 0.855) or the proportion of patients who experienced side-effects [15.6% (88/560) versus 15.3% (93/607), p = 0.881; Supplemental Table 7]. When the proportion of patients who were taking incongruent medications at baseline was evaluated separately, those who switched to congruent medications by week 8 had a significantly lower mean number of side effects compared to those who remained incongruent (0.065 versus 0.242, p = 0.002; Supplemental Table 7). Significantly fewer patients who switched to congruent medications experienced side effects compared to those who did not [6.5% (5/77) versus 16.5% (22/136), p = 0.045].

4. Discussion

Data from this GUIDED trial help clarify a number of concerns and applications regarding the utility of pharmacogenomic testing in patients with difficult-to-treat MDD. The primary outcome of symptom improvement was not significantly different between the study arms. While it is possible this result illustrates real-world clinical outcomes, it may also reflect the fact that three-quarters of patients in TAU were
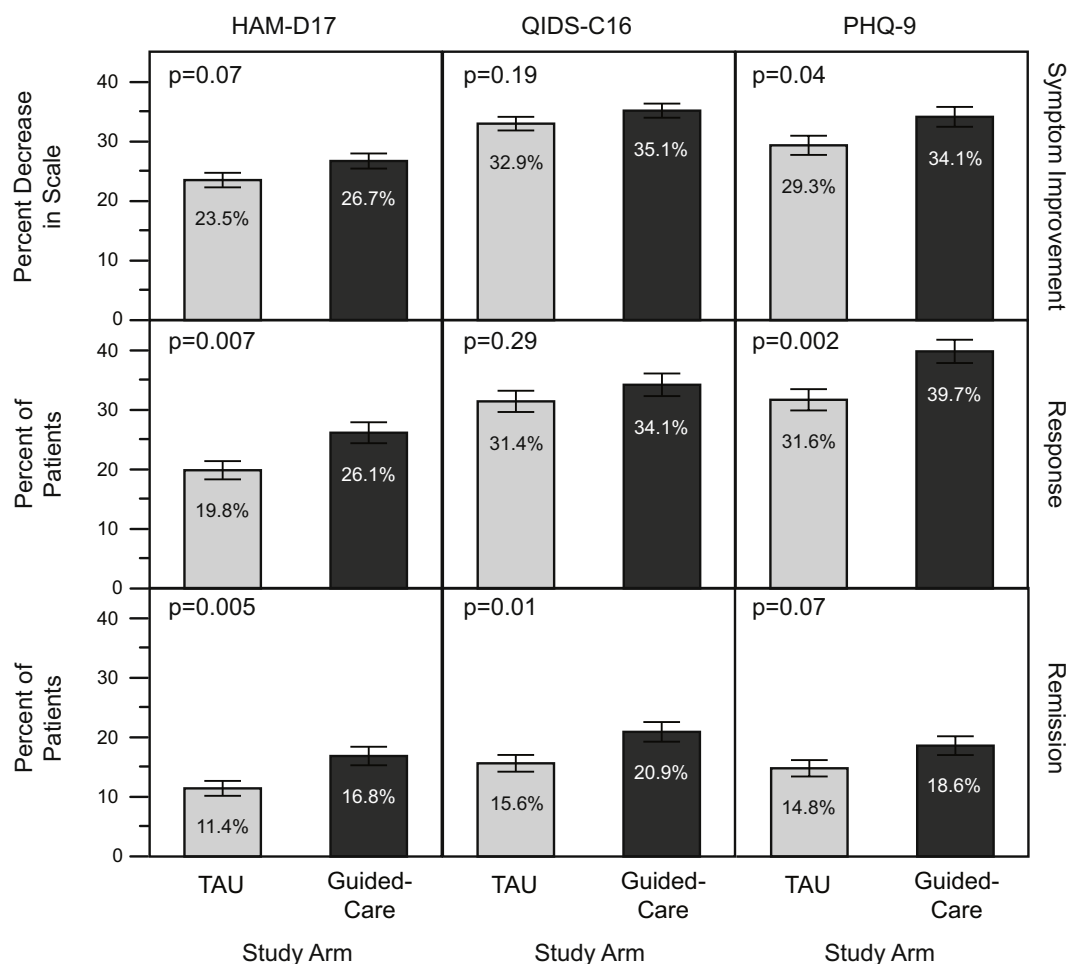
Fig. 2. Patient outcomes in the intent-to-treat cohort at week 8 (n = 1299). Outcomes were evaluated using the HAM-D17, QIDS- C16, and PHQ-9 depression rating scales.
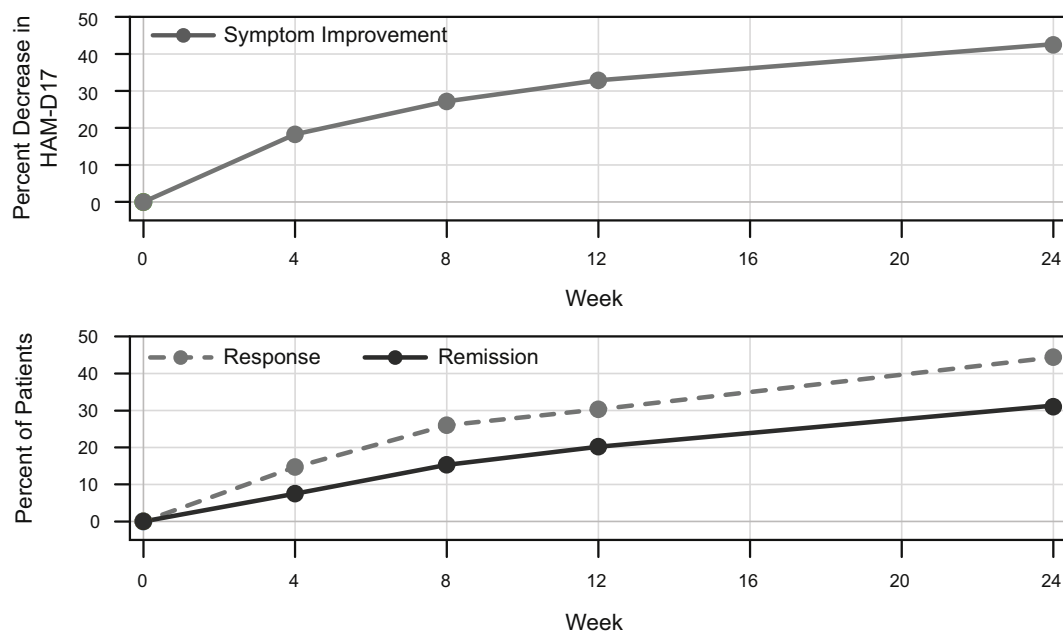

Fig. 3. Durability of improvements in patient outcome throughout the 24-week study in the pharmacogenomics guided-care arm. Outcomes were evaluated using the HAM-D17 depression rating scale.
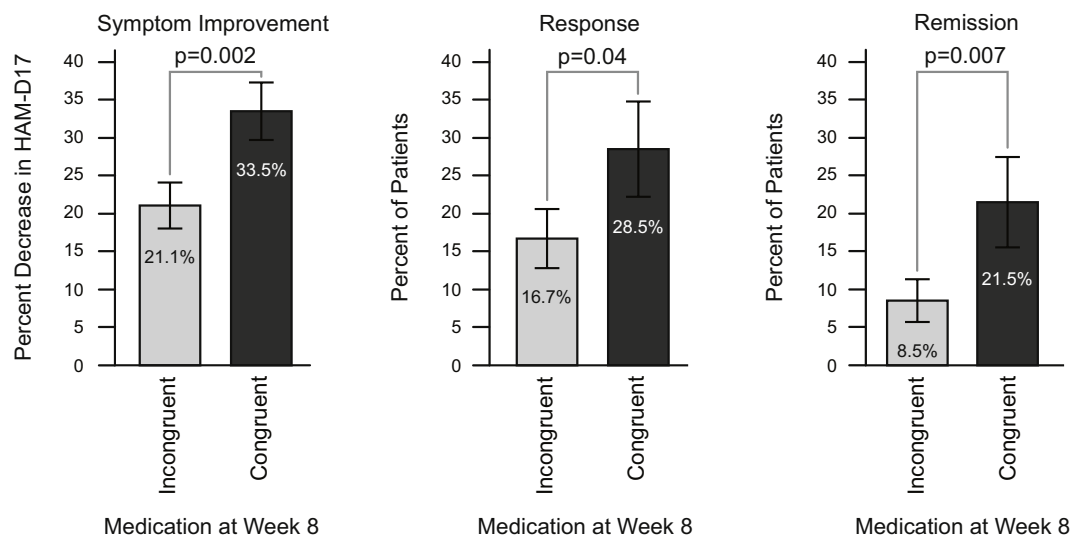
Fig. 4. Patient outcomes among those who were taking incongruent medications at baseline in both study arms (n = 213). Patients were evaluated according to whether they were prescribed congruent (n = 77) or incongruent (n = 136) medications at week 8. Outcomes were evaluated using the HAM-D17 depression rating scale.

incidentally prescribed medications that were congruent with the test report at baseline and throughout the study. Because a key potential clinical application of pharmacogenomics is to identify patients whose treatment resistance may be linked to genetically incongruent medications (Hall-Flavin et al., 2013; Winner, J.G. et al., 2013; Winner et al., 2015), the overall impact of pharmacogenomic testing in this trial may have been diluted by the large proportion of patients already taking genetically congruent medications. When only patients taking genetically incongruent medications at baseline were assessed, symptom improvement was significantly better among patients who switched to congruent medications at week 8 compared to those who remained on incongruent medications (Δ = 12.4%, p = 0.002).

Importantly, combined and weighted, multi-gene pharmacogenomic testing for patients with treatment resistance significantly increased the likelihood of achieving response and remission over TAU medication selection strategies. Larger, controlled studies still need to be performed for those receiving their initial treatment. The modest rates of response and remission reported here are consistent with previous reports of patients with treatment resistant depression (Rush et al., 2006b) and again highlight the significant clinical challenges in this difficult-to-treat population. As such, the modest but important improvements in response and remission for patients in the guided-care arm are clinically meaningful. In addition, *sustained* remission is the ideal objective of treatment for those with MDD (Armstrong, 2011; Cleare et al., 2015; Lam et al., 2016; Rush et al., 2006a); however, most studies on pharmacogenomics have ended after 8 weeks – the typical duration of acute phase treatment. In this trial, the rate of remission doubled from week 8 to week 24 among patients in the guided-care arm. These data suggest that improved patient outcomes achieved with pharmacogenomic testing are durable in the maintenance therapy setting.

The pharmacogenomic test used in this study evaluated gene-drug interactions that may impact medication safety. As such, changes from incongruent to congruent medications should reduce side effect burden. There were no significant differences in side effect burden in the full study cohort. In addition, not all gene-drug interactions were predicted to impact drug safety. However, when the subset of patients taking incongruent medications at baseline were evaluated, side effect burden was significantly reduced when patients switched to congruent medications.

The existing and growing array of pharmacogenomic tests poses clinical implementation challenges. There are several approaches to pharmacogenomic testing and the clinical utility of each approach must

be subject to individual (Bousman and Dunlop, 2018) and Center for Disease Control assessments of robust validation (Prevention, 2010). Previous studies have shown that single-gene testing is often ineffective at improving patient outcomes (EGAPP Working Group, 2007) while studies on multi-gene testing have been mixed. However, several small (< 325 patients) randomized controlled trials have shown that antidepressant selection guided by multi-gene pharmacogenomic testing improved outcomes among patients with MDD (Bradley et al., 2018; Perez et al., 2017; Singh, 2015). The largest effect sizes reported to date emerged from a post-hoc analysis of the subset of patients with severe depression (n = 93) in a randomized controlled trial, though no results were reported for the full study (Bradley et al., 2018).

The clinical utility of the pharmacogenomic test used in this report has been evaluated in three previous studies (Hall-Flavin et al., 2013; Hall-Flavin et al., 2012; Winner, J.G. et al., 2013). A small, pilot, blinded, randomized controlled trial provided the basis for the design of the current study (Winner, J.G. et al., 2013). While underpowered to detect true differences between groups, outcomes improved in the direction of the guided-care group (Winner, J.G. et al., 2013). In addition, treatment guided by pharmacogenomics resulted in significantly better symptom improvement in two open-label, non-randomized trials (Hall-Flavin et al., 2012, 2013). The current randomized, controlled trial demonstrated the clinical utility of pharmacogenomic testing in 1167 patients with MDD whose depression ranged from moderate to very severe. While this study adds to the body of evidence on combined and weighted, multi-gene pharmacogenomics, the conclusions from this trial cannot be generalized to the entire array of pharmacogenomic tests.

This study had several strengths. First, the study design is in line with the recent FDA draft guidance for MDD trials, which recommends evaluation of outcomes for patients with treatment resistant depression ($\geq$1 medication failure) in the short-term (6–8 weeks) and maintenance ($\geq$6 months) settings using HAM-D17 (U.S. Food and Drug Administration, 2018). Second, the diversity of the study cohort mirrors the varied clinical settings in which MDD is treated. Patients were included from both academic and community sites and treating clinicians included psychiatrists and primary care providers. Third, the active treatment approach employed in both study arms models clinical practice and provides a meaningful evaluation of clinical utility. This also represents a significant hurdle in demonstrating efficacy due to the expected range of response to antidepressants in both arms. A systematic review of phase III antidepressant approval studies over 20

years, where active treatment was compared to placebo, shows that response and remission were significantly improved in only 30% and 13% of trials, respectively (Supplemental Table 8). Pharmacogenomic-guided care surmounted the challenge of a comparison to an active control condition and achieved greater response and remission rates in this trial.

There were limitations of this study. First, the treating clinician was not blinded to study arm. This was necessitated by the ethical issues of mandating prescribed medications in order to blind clinicians. To mitigate this limitation, central raters, site raters, and patients were blinded to study arm until after week 8. In addition, the primary assessment was performed by a blinded-central rater who had no interaction with the clinician. Second, the majority of the cohort was Caucasian. While this may limit the generalizability of these findings, the ancestry distribution reported here mimics the US population (US Census Bureau, 2016). Additional evaluation of more diverse populations may be warranted. Third, the per-protocol cohort only included patients with moderate to severe MDD, based on HAM-D17 score. While the intent-to-treat cohort included patients with more mild depression according to HAM-D17 score, the primary results may not be generalizable for patients with mild depression. Finally, the impact of polypharmacy on patient outcomes was not evaluated and is an area for future study. However, any confounding effects of polypharmacy would likely be biased towards the null hypothesis, as medications taken to treat non-MDD conditions would impact patient outcomes in both arms. Despite these potential confounding effects, improved patient outcomes were observed in the guided-care arm over TAU.

In summary, this randomized controlled trial found that weighted and combined multi-gene pharmacogenomic testing significantly increased clinical response and remission rates for patients with MDD in the guided-care arm versus TAU. Pharmacogenomic testing predominantly helped those patients whose treatment resistance may have been related to genetically incongruent medications. Without testing, patients and clinicians are unaware of potential ongoing gene-drug interactions. These results from the GUIDED trial indicate that pharmacogenomic testing is effective in improving response and remission rates among those with prior treatment resistance, particularly for patients who are treated with medications that are incongruent with their genetic profile.

Disclosure

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpsychires.2019.01.003.

References

Altar, C.A., Carhart, J.M., Allen, J.D., Hall-Flavin, D.K., Dechairo, B.M., Winner, J.G., 2015. Clinical validity: combinatorial pharmacogenomics predicts antidepressant responses and healthcare utilizations better than single gene phenotypes. Pharmacogenomics J. 15 (5), 443–451.

Armstrong, C., 2011. APA releases guideline on treatment of patients with major depressive disorder. Am. Fam. Physician 83 (10), 1219–1227.

Bousman, C.A., Dunlop, B.W., 2018. Genotype, phenotype, and medication recommendation agreement among commercial pharmacogenetic-based decision support tools. Pharmacogenomics J Epub ahead of print. https://doi.org/10.1038/s41397-41018-40027-41393.

Bradley, P., Shiekh, M., Mehra, V., Vrbicky, K., Layle, S., Olson, M.C., Maciel, A., Cullors, A., Garces, J.A., Lukowiak, A.A., 2018. Improved efficacy with targeted pharmacogenetic-guided treatment of patients with depression and anxiety: a randomized clinical trial demonstrating clinical utility. J. Psychiatr. Res. 96, 100–107.

Brown, L.C., Lorenz, R.A., Li, J., Dechairo, B.M., 2017. Economic utility: combinatorial pharmacogenomics and medication cost savings for mental health care in a primary care setting. Clin. Therapeut. 39 (3), 592–602 e591.

Cleare, A., Pariante, C.M., Young, A.H., Anderson, I.M., Christmas, D., Cowen, P.J., Dickens, C., Ferrier, I.N., Geddes, J., Gilbody, S., Haddad, P.M., Katona, C., Lewis, G., Malizia, A., McAllister-Williams, R.H., Ramchandani, P., Scott, J., Taylor, D., Uher, R., 2015. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2008 British Association for Psychopharmacology guidelines. J. Psychopharmacol. 29 (5), 459–525.

EGAPP Working Group, 2007. Recommendations from the EGAPP Working Group: testing for cytochrome P450 polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. Genet. Med. 9 (12), 819–825.

Ghio, L., Gotelli, S., Marcenaro, M., Amore, M., Natta, W., 2014. Duration of untreated illness and outcomes in unipolar depression: a systematic review and meta-analysis. J. Affect. Disord. 152–154, 45–51.

Goldberg, J.F., 2017. Do you order pharmacogenetic testing? Why? J. Clin. Psychiatr. 78 (8), 1155–1156.

Greden, J.F., 2013. Workplace depression: personalize, partner, or pay the price. Am. J. Psychiatry 170 (6), 578–581.

Hall-Flavin, D.K., Winner, J.G., Allen, J.D., Jordan, J.J., Nesheim, R.S., Snyder, K.A., Drews, M.S., Eisterhold, L.L., Biernacka, J.M., Mrazek, D.A., 2012. Using a pharmacogenomic algorithm to guide the treatment of depression. Transl. Psychiatry 2, e172.

Hall-Flavin, D.K., Winner, J.G., Allen, J.D., Carhart, J.M., Proctor, B., Snyder, K.A., Drews, M.S., Eisterhold, L.L., Geske, J., Mrazek, D.A., 2013. Utility of integrated pharmacogenomic testing to support the treatment of major depressive disorder in a psychiatric outpatient setting. Pharmacogenetics Genom. 23 (10), 535–548.

Hicks, J.K., Bishop, J.R., Sangkuhl, K., Muller, D.J., Ji, Y., Leckband, S.G., Leeder, J.S., Graham, R.L., Chiulli, D.L., A, L.L., Skaar, T.C., Scott, S.A., Stingl, J.C., Klein, T.E., Caudle, K.E., Gaedigk, A., 2015. Clinical pharmacogenetics implementation consortium (CPIC) guideline for CYP2D6 and CYP2C19 genotypes and dosing of selective serotonin reuptake inhibitors. Clin. Pharmacol. Ther. 98 (2), 127–134.

Hornberger, J., Li, Q., Quinn, B., 2015. Cost-effectiveness of combinatorial pharmacogenomic testing for treatment-resistant major depressive disorder patients. Am. J. Manag. Care 21 (6), e357–365.

Jablonski, M.R., King, N., Wang, Y., Winner, J.G., Watterson, L.R., Gunselman, S., Dechairo, B.M., 2018. Analytical validation of a psychiatric pharmacogenomic test. Pers. Med. 15 (3), 189–197.

Lam, R.W., McIntosh, D., Wang, J., Enns, M.W., Kolivakis, T., Michalak, E.E., Sareen, J., Song, W.Y., Kennedy, S.H., MacQueen, G.M., Milev, R.V., Parikh, S.V., Ravindran, A.V., 2016. Canadian Network for Mood and anxiety treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: section 1. Disease burden and principles of care. Can. J. Psychiatr. 61 (9), 510–523.

Maciel, A., Cullors, A., Lukowiak, A.A., Garces, J., 2018. Estimating cost savings of pharmacogenetic testing for depression in real-world clinical settings. Neuropsychiatric Dis. Treat. 14, 225–230.

Mrazek, D.A., Hornberger, J.C., Altar, C.A., Degtiar, I., 2014. A review of the clinical, economic, and societal burden of treatment-resistant depression: 1996-2013. Psychiatr. Serv. 65 (8), 977–987.

Organization, W.H., 2017. Depression fact sheet. 2018. http://www.who.int/mediacentre/factsheets/fs369/en/, Accessed date: 28 March 2018.

Perez, V., Salavert, A., Espadaler, J., Tuson, M., Saiz-Ruiz, J., Saez-Navarro, C., Bobes, J., Baca-Garcia, E., Vieta, E., Olivares, J.M., Rodriguez-Jimenez, R., Villagran, J.M., Gascon, J., Canete-Crespillo, J., Sole, M., Saiz, P.A., Ibanez, A., de Diego-Adelino, J., Menchon, J.M., 2017. Efficacy of prospective pharmacogenetic testing in the treatment of major depressive disorder: results of a randomized, double-blind clinical trial. BMC Psychiatry 17 (1), 250.

Prevention, C.f.D.C.a., 2010. ACCE Model List of 44 Targeted Questions Aimed at a Comprehensive Review of Genetic Testing.

Rosenblat, J.D., Lee, Y., McIntyre, R.S., 2017. Does pharmacogenomic testing improve clinical outcomes for major depressive disorder? A systematic review of clinical trials and cost-effectiveness studies. J. Clin. Psychiatr. 78 (6), 720–729.

Rush, A.J., Kraemer, H.C., Sackeim, H.A., Fava, M., Trivedi, M.H., Frank, E., Ninan, P.T., Thase, M.E., Gelenberg, A.J., Kupfer, D.J., Regier, D.A., Rosenbaum, J.F., Ray, O., Schatzberg, A.F., 2006a. Report by the ACNP Task Force on response and remission in major depressive disorder. Neuropsychopharmacology 31 (9), 1841–1853.

Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Niederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., McGrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., Fava, M., 2006b. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am. J. Psychiatry 163 (11), 1905–1917.

Singh, A.B., 2015. Improved antidepressant remission in major depression via a pharmacokinetic pathway polygene pharmacogenetic report. Clin Psychopharmacol Neurosci 13 (2), 150–156.

Swen, J.J., Nijenhuis, M., van Rhenen, M., de Boer-Veger, N.J., Buunk, A.M., Houwink, E.J.F., Mulder, H., Rongen, G.A., van Schaik, R.H.N., van der Weide, J., Wilffert, B., Deneer, V.H.M., Guchelaar, H.J., 2018. Pharmacogenetic information in clinical guidelines - the European perspective. Clin. Pharmacol. Ther. 103 (5), 795–801.

Trivedi, M.H., Morris, D.W., Wisniewski, S.R., Lesser, I., Nierenberg, A.A., Daly, E., Kurian, B.T., Gaynes, B.N., Balasubramani, G.K., Rush, A.J., 2013. Increase in work productivity of depressed individuals with improvement in depressive symptom severity. Am. J. Psychiatry 170 (6), 633–641.

US Census Bureau, 2016. In: Bureau, U.S.C. (Ed.), ACS Demographic and Housing Estimates: 2012-2016 American Community Survey 5-Year Estimates.

U.S. Food & Drug Administration, 2018. Major depressive disorder: developing drugs for treatment - guidance for industry (draft guidance). In: U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). Food and Drug Administration. www.FDA.gov.

Winner, J., Allen, J.D., Altar, C.A., Spahic-Mihajlovic, A., 2013a. Psychiatric pharmacogenomics predicts health resource utilization of outpatients with anxiety and depression. Transl. Psychiatry 3, e242.

Winner, J.G., Carhart, J.M., Altar, C.A., Allen, J.D., Dechairo, B., 2013b. A prospective, randomized, double-blind study assessing the clinical impact of integrated pharmacogenomic testing for major depressive disorder. Discov. Med. 16, 219–227.

Winner, J.G., Carhart, J.M., Altar, C.A., Goldfarb, S., Allen, J.D., Lavezzari, G., Parsons, K.K., Marshak, A.G., Garavaglia, S., Dechairo, B.M., 2015. Combinatorial pharmacogenomic guidance for psychiatric medications reduces overall pharmacy costs in a 1 year prospective evaluation. Curr. Med. Res. Opin. 31 (9), 1633–1643.

Zeier, Z., Carpenter, L.L., Kalin, N.H., Rodriguez, C.I., McDonald, W.M., Widge, A.S., Nemeroff, C.B., 2018. Clinical implementation of pharmacogenetic decision support tools for antidepressant drug prescribing. Am. J. Psychiatry 175 (9), 873–886 appiajp.201817111282.

Zhou, Z.W., Chen, X.W., Sneed, K.B., Yang, Y.X., Zhang, X., He, Z.X., Chow, K., Yang, T., Duan, W., Zhou, S.F., 2015. Clinical association between pharmacogenomics and adverse drug reactions. Drugs 75 (6), 589–631.

Zubenko, G.S., Sommer, B.R., Cohen, B.M., 2018. On the marketing and use of pharmacogenetic tests for psychiatric treatment. JAMA Psychiatry 75 (8), 769–770.