**Pharmacogenomic Testing and Personalized Treatment of Depression**

Roy H. Perlis*

**BACKGROUND:** There is wide variation in antidepressant efficacy and tolerability during the treatment of major depressive disorder, a brain disease associated with significant morbidity and mortality risk. The ability to rapidly identify optimal treatment, thereby shortening the time to symptomatic remission, could reduce these risks and associated costs.

**CONTENT:** Up to 42% of variance in antidepressant response is associated with common genetic variation, and there are over 10 psychotropic medications for which the US Food and Drug Administration–approved labeling reflects a genetic test. Most published studies have examined functional variations in genes of the cytochrome p450 system, relevant to metabolism of many antidepressants. However, there are few data supporting the clinical usefulness of specific pharmacogenetic tests. Randomized trials and cost-effectiveness studies are emerging, but larger-scale studies are needed. Specific challenges in translating genetic association results to clinical practice include need for replication to address risk of type I error, overestimation of effect sizes, absence of data from generalizable cohorts, and absence of comparative data that would suggest one specific intervention over another. Several opportunities to accelerate development and validation of new tools for stratification remain, including integration of these tests with clinical data or other biomarkers and application of electronic health records for test development and investigation.

**SUMMARY:** Although common genetic variation, particularly in genes of the cytochrome p450 system, has been associated with antidepressant response, evidence that this variation may be successfully applied to guide treatment selection is just emerging. Larger-scale studies facilitated by informatics tools will clarify the usefulness of such tests.

© 2013 American Association for Clinical Chemistry

Only about one-third of individuals with major depressive disorder (MDD) achieve symptomatic remission with an initial antidepressant treatment, and rates of attrition and nonadherence during antidepressant treatment are high. Prolonged depressive episodes are associated with increased impact on quality of life and functioning, risk for suicide, and greater societal costs. The ability to more rapidly identify an optimal depression treatment, pharmacologic or otherwise, could shorten the time to effective treatment and reduce these costs.

Many of the brain disorders treated in psychiatry are strongly heritable. Estimating the heritability of treatment response to psychotropic medications using traditional twin-based approaches is more challenging, because of the proliferation of treatment options and limited reliability of retrospective assessment. On the basis of genome-wide association study (GWAS) data, however, the proportion of variance in antidepressant response explained by common genetic variation has been estimated to be as much as 42%.

The US Food and Drug Administration has already incorporated labeling information related to genetic testing results for more than 100 medications, including at least 9 antidepressants and multiple other medications used together with antidepressants. Conversely, reviews of the evidence base for pharmacogenomic testing in antidepressant response emphasize the paucity of evidence supporting their clinical application.

To better understand this apparent discordance, this review begins by considering published genetic association data for antidepressant efficacy and tolerability. It then summarizes the small number of random-
ized trials and cost-effectiveness studies examining pharmacogenomics testing in treatment of depression. Finally, it concludes with caveats about the published data and directions for future research, with an eye toward clinical application of biomarkers as prediction tools.

**State of the Field: Association Studies**

**PHARMACOKINETIC VARIATION**

Attempts to understand the impact of common genetic variation on treatment responses in MDD have focused on the cytochrome p450 (CYP450) hepatic enzymes. Among these, CYP450 2D6 is responsible for oxidative metabolism of up to 25% of medications (5), including a large subset of antidepressants. Relatively common functional variations in the genes coding for 2D6 and other CYP450 enzymes have been described. For purposes of analysis, these functional variations are aggregated to derive an overall metabolic phenotype for each CYP450 enzyme. For 2D6, the most commonly observed phenotype is referred to as an extensive metabolizer (EM), and is observed in 60%–85% of white individuals. Individuals with 2 disrupted copies of this gene are referred to as poor metabolizers (PMs) and would be expected to have greater-than-typical blood concentrations from a given medication dosage. Those individuals with 1 disruption, or 2 variants with modestly diminished function, are considered intermediate metabolizers (IMs). Individuals with a duplication of 1 of these genes may metabolize substrates more efficiently, and are referred to as ultrarapid metabolizers (UMs). For all of these categories, marked variations between ethnic groups have been observed, in terms of prevalence of functional groups as well as specific alleles observed. For example, 2D6 UMs are observed more frequently in populations of Southern European, Saudi Arabian, and Ethiopian origin (5).

The role of CYP450 variation in tricyclic antidepressant metabolism has recently been reviewed as part of a guidance document from a pharmacogenomics research consortium (6). For the relatively small subset of tricyclic antidepressants studied, the pharmacokinetic studies reviewed clearly document effects of these metabolic categories on blood concentrations.

However, tricyclic antidepressants are increasingly rare in the management of depression because of their narrow therapeutic index, which also makes an optimal blood concentration particularly important. First line treatment of MDD now relies on selective serotonin reuptake inhibitors (SSRIs). A comprehensive review of CYP450 studies in 2007 identified a total of 37 studies, nearly all in white cohorts, and concluded that the evidence of association with treatment outcomes was “mixed” based upon “heterogeneous studies in small samples” (7). On the other hand, a previous systematic review indicated 2D6 was relevant in nearly half of the antidepressants examined, and proposed dosage adjustments accordingly (8). In the most recent published guideline, recommended dose adjustment was confined to 5 of the newer antidepressants (venlafaxine, paroxetine, citalopram, escitalopram, and sertraline) (9).

Even in cases in which CYP450 status can be demonstrated to influence metabolism of a given drug, there are 2 additional conditions that must be satisfied to establish that testing for genetic variation could be useful. First, the effect must be substantial enough to meaningfully impact blood concentrations even after accounting for the many nongenetic factors that influence drug metabolism, e.g., smoking status, diet, concomitant medications, and adherence. In particular, the low rate of adherence to antidepressant treatment in clinical settings and its impact on outcomes has been well established. Because all of these factors change over time, the only direct means of characterizing a patient’s blood concentration for a given antidepressant is to measure it directly, which can be done in many clinical laboratories.

In a striking demonstration of the challenge of relying on genetic data alone, 900 individuals with depression treated with the serotonin–norepinephrine reuptake inhibitor (SNRI) venlafaxine were characterized in terms of CYP450 2D6 status (10). Venlafaxine is a “prodrug” converted to the active drug O-desmethylvenlafaxine (desvenlafaxine) by this enzyme, so serum concentrations of both venlafaxine and desvenlafaxine were measured and compared to predicted metabolic status on the basis of 2D6 genotyping. Although 4% of individuals were predicted to be PMs based on genotype, the ratio of desvenlafaxine to venlafaxine indicated that 27% were functionally PMs, indicating the critical role played by environmental effects.

The second condition that must be satisfied is that the resulting blood concentrations must also be shown to influence outcomes. Although it may seem counterintuitive, there is no convincing evidence of a dose response or therapeutic window with SSRIs and SNRIs, with the possible exception of venlafaxine [reviewed in (11); a small number of studies do support dose increase as a strategy in initial treatment nonresponders (12)]. At the extremes, blood concentrations likely matter a great deal. Individuals with very low concentrations are unlikely to respond, whereas those with very high concentrations may be more apt to experience adverse effects. However, within this range the existence of optimal or therapeutic concentrations has not been established.

To date, the strongest evidence for CYP450 effects among newer antidepressants comes from studies of venlafaxine. Individuals who metabolize venlafaxine less efficiently will have lower blood concentrations of...
the active metabolite desvenlafaxine. As hypothesized, in a post hoc analysis of 4 venlafaxine clinical trials in MDD, outcomes among PMs (defined by blood concentrations, not genotype) were poorer than those of extensive metabolizers, with fewer individuals reaching symptomatic remission (13). Importantly, however, because this study measured metabolic status directly from drug concentrations, it may have overestimated the importance of genetic contributions to outcomes. Moreover, this example, in which the administered drug is actually a prodrug requiring activation by metabolism, is relatively unusual among psychotropics.

The clinical implication of CYP450 status for other antidepressants is less clear. Two small studies suggest CYP450 2D6 PMs are more likely to experience adverse effects. Among 18 antidepressant-treated patients presenting with adverse effects, 44% were PMs, compared to a rate of 21% observed among 56 other depressed patients (14). In another pilot study of 45 psychiatric inpatients treated with cytochrome CYP450 2D6 substrates, a greater number of adverse effects were observed in PMs compared to EMs and UMs (15).

Conversely, CYP450 2D6 UMs might be expected to demonstrate decrements in efficacy rather than tolerability. Indeed, a pilot investigation found that 8 of 81 individuals (10%) who had failed to respond to a CYP450 2D6 substrate antidepressant, including tricyclic antidepressants but also newer agents, were UMs, a markedly higher prevalence than anticipated in Northern European populations (16). Although these data are promising, they must be interpreted cautiously in light of the small sample sizes. The impact of PM status at CYP450 2C19 has also been investigated in 1 large cohort. Among 1074 citalopram-treated patients, PMs tolerated citalopram more poorly than other patients (17).

Although pharmacokinetic investigations have focused on CYP450 systems, the drug transport protein P-glycoprotein [ATP-binding cassette, sub-family B (MDR/TAP), member 1 (ABCB1)]3 has also been investigated, on the basis of evidence of its role in the efflux of medications across the blood– brain barrier (18). A common variant was associated with escitalopram dose required for remission in a study of 57 patients, although the greatest dose was required for heterozygous individuals, a result difficult to explain in biological terms (19). The same study also suggested that this variant was associated with greater venlafaxine remission rates but not effective dosage, illustrating the challenges inherent in interpreting studies with multiple phenotypic outcomes. A large (n = 1257) incident user cohort indicated that a particular haplogroup was associated with greater likelihood of subsequent switch to a different antidepressant (20). On the other hand, multiple antidepressant response cohorts did not support an association between common variants in ABCB1 and overall treatment efficacy (21).

Most recently, variation in this gene has also been investigated in a longitudinal study; 6 single-nucleotide polymorphisms genotyped in 424 individuals with major depression were examined for association with adverse drug effects. Two single-nucleotide polymorphisms were associated with greater numbers of adverse effects overall, with particular effects noted for insomnia, sexual adverse effects, and gastrointestinal symptoms (22).

**PHARMACODYNAMIC VARIANTS**

Spanning more than a decade, a plethora of candidate gene studies investigated targets related to the monoaminergic hypothesis of antidepressant effect, and serotonergic neurotransmission in particular. The best studied of these was the serotonin transporter [solute carrier family 6 (neurotransmitter transporter), member 4 (SLC6A4)], the primary site of action of SSRIs, which has been associated to everything from amygdala activation to risk for depression following stressful life events to interpretive dance. Recent meta-analyses generally support a modest association with SSRI efficacy in whites but not in Asians (odds ratio for remission, 1.53; 95% CI, 1.16–2.16) (23). A major caveat is the risk of publication bias, which was in fact detected in that study. The specificity of effect, if any, has also not been established; 1 study in Korean in- and outpatients (n = 241) suggested association with response to both SSRI and the norepinephrine reuptake inhibitor nortriptyline. Some early, smaller studies suggested that the apparent association with efficacy could be mediated by differences in tolerability (24), but metaanalysis of 17 studies, including 2504 white Europeans, found no association with treatment discontinuation (25).

Interpretation of other candidate genes is also challenging because of inconsistent results. Although a host of genes have been associated with treatment response in individual cohorts, including the serotonin 2A receptor gene [5-hydroxytryptamine (serotonin) receptor 2A, G protein-coupled (HTR2A)], metaanalysis among large cohorts discussed below does not provide strong support for any of them to date (21).

**GWAS RESULTS**

As in medical genetics as a whole, the focus of investigation of genetics in depression has shifted to less biased, genome-wide investigations. Two recent, partially overlapping metaanalyses of antidepressant response failed to
identify any genome-wide associations with efficacy of an initial antidepressant trial. In the first, the 3 largest published antidepressant response cohorts, drawn from the UK, Germany, and the US, were harmonized and metaanalyzed (21). The second metaanalysis, Newmeds, added additional cohorts to the previously studied UK cohort (26). Most disappointingly, a “polygenic” analysis examining the overlap in liability between Newmeds and the multicenter Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study found little evidence of shared risk. On the other hand, the failure to detect genome-wide associations is perhaps not surprising in light of the apparent complexity of the phenotype and the limited cohort size, which is just approaching the scale at which initial psychiatric disease associations were identified.

Other pharmacogenomic analyses using more extreme phenotypes, such as statin myopathy, have succeeded despite modest cohort sizes. A recent antidepressant study adopted a similar approach, examining the more extreme and potentially more clinically actionable phenotype of treatment resistance (27). That study contrasted initial treatment responders with those individuals who did not respond despite at least 2 antidepressant interventions, and included a cohort of individuals identified from electronic health records. However, to date, only analyses of rare genetic variation, rather than GWAS results, have been reported.

A far more common antidepressant adverse effect that contributes to treatment discontinuation is sexual dysfunction. In one of the only genome-wide association studies of antidepressant outcome in a nonwhite population, investigation of a small Japanese cohort (n = 201) treated with SSRIs or SNRIs did not identify any loci associated with sexual dysfunction at a conventional genome-wide threshold (28). Previous investigations of candidate neurotransmitter genes provided modest support for involvement of glutamatergic genes (29), though without subsequent replication.

Although beyond the scope of this review, atypical antipsychotics are increasingly used in the treatment of MDD. The risk of adverse effects such as weight gain and emergence of diabetes mellitus can be substantial and has motivated recent GWAS. For example, a study of antipsychotic-associated weight gain in a treatment-naive pediatric cohort implicated a common variant in a melanocortin receptor, MC4R (30), although this result awaits replication and clarification of treatment specificity.

**Evidence for Efficacy of Testing to Improve Outcomes**

A 2007 review found no studies supporting the efficacy or utility of CYP450 testing for antidepressant response (7). Since that time, a small number of randomized, controlled trials have directly investigated the impact of pharmacogenomics testing on antidepressant efficacy. In the only investigation with a partially blinded design, 51 adult outpatients with MDD were randomized 1:1 to pharmacogenetic testing or treatment as usual and monitored biweekly for 8 weeks, using a panel that included CYP450 2D6, 2C19, and 1A2 as well as 2 serotonergic variants previously suggested to be associated with antidepressant response. Patients and clinical raters, but not clinicians, were blinded to test status. Although numerically the results suggested a modestly greater reduction in depression severity at all time points, this improvement reached statistical significance only at week 4. After 8 weeks, 31% improvement was observed in the tested group vs 19% in the treatment-as-usual group (P = 0.3). Specific interventions and genotypes were not examined, although clinician survey suggested that the results were primarily applied to select an initial treatment or to increase confidence in that treatment.

Two similar open-label, nonrandomized, 8-week studies of outpatients with MDD using the same commercial panel yielded concordant results. In a smaller cohort (n = 22 patients with treatment as usual, followed by 22 with assay-guided treatment) mean improvement in depression severity as measured by the Hamilton Depression Rating Scale (HAMD-17) was 31% among assay-guided patients vs 18% among the treatment-as-usual group (P = 0.04). In a larger study of similar design, 114 individuals were compared to 113 in a treatment-as-usual cohort (31). A 47% reduction in HAMD-17 score in the guided group was observed at week 8 vs a 30% reduction among the unguided group (P < 0.0001).

A key limitation of the latter 2 studies is the lack of blinding of patients or clinicians, a particular concern in light of the substantial rate of placebo response to antidepressant treatment. A secondary analysis of these 2 cohorts suggested greatest relative benefit among individuals treated with drugs more likely to require caution or monitoring at baseline, arguing that at least some of the observed difference does not arise solely from expectancy. A second limitation is that these studies investigated aggregate results, combining CYP450 testing with pharmacodynamic results for which the evidence of informativeness may be less strong. Thus, the relative contribution of each aspect of the test cannot be identified. An important possibility to examine is the extent of benefit associated with CYP450 testing, which will be useful in guiding future test analysis and design. A final caveat in interpreting the 2 latter studies is that they were not randomized, but rather enrolled a treatment-as-usual cohort followed by an assay-guided treatment, leading to groups
that were not entirely balanced. For example, in the larger investigation, the treatment-as-usual group had been treated on average with 1 additional psychiatric medication before entering the study (31).

Evidence for Utility or Cost-Effectiveness

An intervention may be efficacious without being cost-effective. Initial attempts to model the cost-effectiveness of an antidepressant pharmacogenomics test, using outcomes from the multicenter STAR*D effectiveness study, suggested that a test in treatment-naïve patients would likely not meet an accepted threshold for cost-effectiveness (32), but that under certain conditions (i.e., larger effect sizes) this threshold could be met. However, these in silico models must make many assumptions about costs and test parameters and may be most useful in planning prospective investigations.

Utilization data provide some support for the potential cost-effectiveness of CYP450 testing in specific psychiatric populations. In the 45-patient pilot study of inpatients treated with CYP450 substrates noted above, individuals with either poor metabolism or ultrarapid metabolism had 1-year treatment costs approximately $4000–$6000 greater than wild-type metabolizers or IMs, and PMs had longer mean durations of hospitalization (15).

A similar outpatient study examined healthcare utilization among outpatients with depressive or anxiety disorders, but included multiple CYP450 enzymes as well as some pharmacodynamic variants from a commercial panel. In that study, non–wild-type individuals treated with a CYP450 substrate or other “less desirable” drug had greater prior-year costs, including more medical visits and medical absence days and disability claims (33). In both studies, the assumption was that the higher-cost groups represent an opportunity to intervene to reduce costs. However, the observed associations must be considered with the caveat that many of the potential less-preferred interventions are medications used in more severe or treatment-resistant illness (i.e., atypical antipsychotics, tricyclic antidepressants) and thus this finding may represent an example of confounding by indication.

Key Caveats in Evaluating Published Data

In considering the published antidepressant pharmacogenomics data, multiple caveats apply (Fig. 1). Most importantly, the risk of type I error is substantial. Although metaanalysis can help in this regard, the studies included are often quite heterogeneous and require major assumptions about comparability. The problem of publication bias is particularly acute, with small negative studies dismissed as underpowered. This risk is compounded by the complexity of treatment response phenotypes, for which multiple outcomes can be examined in parallel. Even in cases in which such associations are true, there is substantial risk that the magnitude of association is inflated, a phenomenon referred to as “winner’s curse.”

Additional challenges arise in understanding the specificity and generalizability of published results. In nearly all cases, association results pertain to a single treatment. The extent to which a predictor indicates poor outcome independent of treatment, rather than poor outcome with a specific treatment, is critical to clinical application but rarely addressed. Few studies are adequately powered to detect main effects of indi-
individual common genetic variations, much less the variant-by-treatment interactions, which would help to guide treatment selection. A further challenge is generalizability, the problem of extrapolating from homogeneous study populations to heterogeneous clinical cohorts. Most association studies of antidepressant response focus on white Europeans drawn from treatment studies that exclude individuals with substantial psychiatric and medical comorbidity. A metaanalysis suggests that one proposed moderator of antidepressant response might have differential effects between white and Asian populations (23); the prevalence of CYP450 phenotypes is known to differ between ethnic groups as well.

A final consideration is that a potential consequence of testing meriting study is increased use of more expensive interventions with marginal benefit, or avoidance of potentially effective interventions on the basis of misinterpretation of the test. For example, reports that identify CYP450 substrates with red letters or a stop sign for PMs may be misinterpreted as indicating that these drugs are contraindicated (6). Although clinicians are comfortable with reference ranges, the application of more probabilistic test results may require additional education.

Emerging Areas of Investigation

One area of development for personalization of antidepressant treatment may be the rediscovery of clinical risk stratification. The enthusiasm for biomarker identification has overshadowed a long history of efforts to direct prescribing on the basis of clinical observations, i.e., by identifying patient subgroups on the basis of clinical symptomatology or comorbidity (34). For example, anxiety, psychotic-like symptoms, and other sets of depressive symptoms have been associated with poorer treatment response (35–37). Two relatively recent developments might make this strategy more practical: the emergence of larger clinical data sets, including those drawn from large effectiveness studies, electronic health records (38), and patient self-report in direct-to-consumer genomics testing (39), as well as the application of newer machine learning tools. However, although there are numerous reports of putative predictors, there have been very few efforts to translate even clinical predictors to practice (40).

The opportunity to integrate clinical and genetic data for more precise prediction of antidepressant outcomes also remains to be explored. In other areas such as cardiovascular disease, combinations of traditional clinical models and emerging genetic data have been shown to modestly improve performance over either type of data alone. For example, I study compared models with clinical predictors of adverse cardiovascular outcomes to models that also incorporated common variants associated with hyperlipidemia. Although overall prediction accuracy was not improved, the capacity to estimate risk in intermediate risk groups (i.e., calibration) was still improved (41).

Nearly all antidepressant pharmacogenomics investigations to date have assessed common variation. Studies have used targeted sequencing of candidate neurotransmitter-related genes to identify novel variants (42), although the functional significance of these variants remains to be established. Whole-genome or whole-exome results have not yet been reported. Whether rare variants meaningfully influence more common phenotypes, as is the case in psychiatric disorders in general, remains to be established (27).

References

6. Hicks JK, Swen JJ, Thorn CF, Sangkuhl K, Kharausch ED, Ellingrod VL, et al. Clinical pharmacogenomics implementation consortium guideline for CYP2D6 and CYP2C19 genotypes and dosing of


