Received: 7 September 2017 | Revised: 4 January 2018 | Accepted: 10 February 2018

DOI: 10.1002/da.22742

RESEARCH ARTICLE

WILEY ADAA ANXIETY AND DEPRESSION ASSOCIATION OF AMERICA

Pharmacogenetic testing among patients with mood and anxiety disorders is associated with decreased utilization and cost: A propensity-score matched study

Roy H. Perlis MD, MSc[1] | Rajesh Mehta RPh, MS[2] | Alison M. Edwards MStat[2] | Arun Tiwari MBA[2] | Guido W. Imbens PhD[3]

[1]Center for Quantitative Health, Division of Clinical Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

[2]Healthagen, New York City, NY, USA

[3]Graduate School of Business, Stanford University, Stanford, CA, USA

Correspondence
Roy H. Perlis, 185 Cambridge Street, 6th Floor, Boston, MA 02114, USA.
Email: rperlis@partners.org

Background: Naturalistic and small randomized trials have suggested that pharmacogenetic testing may improve treatment outcomes in depression, but its cost-effectiveness is not known. There is growing enthusiasm for personalized medicine, relying on genetic variation as a contributor to heterogeneity of treatment effects. We sought to examine the relationship between a commercial pharmacogenetic test for psychotropic medications and 6-month cost of care and utilization in a large commercial health plan.

Methods: We performed a propensity-score matched case-control analysis of longitudinal health claims data from a large US insurer. Individuals with a mood or anxiety disorder diagnosis ($N = 817$) who received genetic testing for pharmacokinetic and pharmacodynamic variation were matched to 2,745 individuals who did not receive such testing. Outcomes included number of outpatient visits, inpatient hospitalizations, emergency room visits, and prescriptions, as well as associated costs over 6 months.

Results: On average, individuals who underwent testing experienced 40% fewer all-cause emergency room visits (mean difference 0.13 visits; $P < 0.0001$) and 58% fewer inpatient all-cause hospitalizations (mean difference 0.10 visits; $P < 0.0001$) than individuals in the control group. The two groups did not differ significantly in number of psychotropic medications prescribed or mood-disorder related hospitalizations. Overall 6-month costs were estimated to be $1,948 (SE 611) lower in the tested group.

Conclusions: Pharmacogenetic testing represents a promising strategy to reduce costs and utilization among patients with mood and anxiety disorders.

KEYWORDS
antidepressant, anxiety disorder, biomarker, cytochrome P450, genetic, major depressive disorder, propensity score

1 | INTRODUCTION

Personalized medicine has become a major focus of modern medicine in the United States of America and internationally, with a particular emphasis on incorporating genetic variation to guide medication prescribing (Institute of Medicine of the National Academies, 2010; Office of the Press Secretary, 2015). The US FDA lists more than 100 drug labels reflecting genetic variation, including multiple boxed warnings, with most addressing variation in a set of genes coding for hepatic cytochrome P450 (CYP450) enzymes known to influence drug levels in vivo (Zhou, Liu, & Chowbay, 2009; Federal Drug Administration, 2016). This list includes 27 medications commonly used in psychiatric practice to treat mood and anxiety disorders. Small cohort studies have suggested that comprehensive assays integrating CYP450 variation with pharmacodynamic variation may be associated with improved global outcomes in depression or in psychiatric outpatients more generally (Brennan et al., 2015).

On the other hand, the potential benefits of genetic testing in real-world clinical practice have not been well characterized in psychiatry. Consistent with analyses of claims data suggesting that CYP450 substrates are associated with greater overall health costs and hospital readmissions (McCoy, Cagan, Roberson, Castro, & Perlis, 2017),

a prior cost-effectiveness analysis using propensity-score matched claims samples suggested that pharmacogenetic testing among psychiatric patients was associated with greater treatment adherence and lower health costs over four months (Fagerness et al., 2014). In order to better understand the impact of testing in psychiatric patients over the longer term, we conducted a substantially larger propensity-score matched case-control analysis drawn from a large commercial insurer, incorporating additional matching variables to mimic as closely as possible a randomized controlled trial in a diverse national health network.

2 | MATERIALS AND METHODS

2.1 | Study design

The study used a propensity-score matched retrospective case-control design in which cases were drawn from a registry of individuals who had received a commercial pharmacogenetic test, the Genecept assay (Genomind Inc.; King of Prussia, PA) between January 1st, 2012 and December 31st, 2015. The control or treatment-as-usual (TAU) cohort was drawn from insurance claims data of a commercial insurer, Aetna, between January 1st, 2012 and December 31st, 2015. Eligible individuals were those aged 18 and older during this period drawn from Aetna's commercially insured or Medicare population including pharmacy benefits, and who were continuously enrolled for 12 months with no break in coverage. In addition, controls were required to be diagnosed with a mood or anxiety disorder (defined as per Supporting Information Table 1) at least twice between July 1st, 2012 and June 30th, 2015, and to have received two or more medication treatment trials as a further means of ensuring comparator patients were not newly diagnosed with a mood disorder. In other words, requiring 2+ failed treatment trials in the control arm was intended to align them more closely to tested cases with respect to disease progression and treatment options, prior to further matching as described below.

Individuals for whom different utilization patterns were anticipated based on prior analyses and reports outside of psychiatry, or for whom pharmacy data were unavailable, were excluded a priori per protocol. This included individuals in a skilled nursing facility, assisted living, or hospice facility; those enrolled in the Aetna Compassionate Care Program for hospice; and those diagnosed with metastatic cancer or dementia during or prior to the study period.

The case, or assay-guided treatment (AGT), group had received a buccal-swab assay integrating variants in ten genes, of which three code for pharmacokinetic variants (including CYP450 2D6, 2C19, and 3A4) and seven code for published pharmacodynamic variants (i.e., receptors, transporters, enzymes, and ion channels) related to common neurotransmitter systems. The net effects of these variants are integrated using an algorithm based on published evidence to indicate groups of treatment options (by class or drug) that may be more or less appropriate. In addition to a written report describing the joint effects of all tested variants, clinicians are encouraged, but not required, to consult by telephone with a pharmacist or physician for additional interpretation. (While other commercial tests were available during the study period, the data available only allowed us to identify tests from one source.) Figure 1 illustrates the process of cohort selection for the experimental (AGT) group, and the control (matching) group. Among 2,735 tested individuals matched to the Aetna Data Warehouse, 2,219 met study inclusion and exclusion criteria; of these, 1,639 had the necessary 6-month period of lead-in and posttest data available.

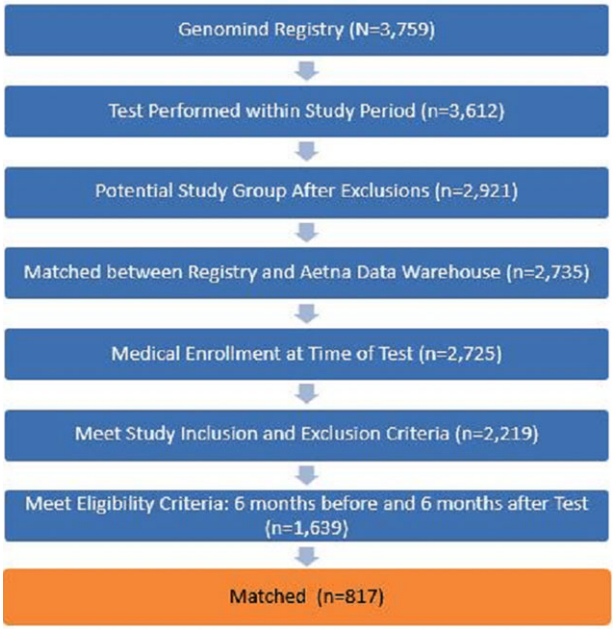

FIGURE 1 CONSORT-like diagram of subject selection

As a retrospective analysis, the study was approved by the Aetna Data Governance (internal) and the Sterling Institutional Review Board (external). Informed consent was not required as no contact with human subjects was involved, and subject identifiers were not released by Aetna. As this was not a clinical trial, it was not registered at clinicaltrials.gov, but the analytic portion of the approved protocol confirming outcome measures is available upon request from the authors.

2.2 | Derivation of baseline and outcome variables

Subject sociodemographic characteristics were extracted from Aetna data, including age, sex, imputed household income based on zip code (applying 2013 US Census bureau data available from the American Community Survey (United States Census Bureau)), insurance type, and funding source for insurance (self-insured versus employer-insured). Baseline clinical characteristics were defined based upon the 6-month lead-in period, including major depression diagnosis, anxiety disorder, or bipolar disorder (Table 1 and Supporting Information Table 2), with diagnoses categorized via ICD-9 codes. Presence of a validated diagnosis of nonpsychiatric comorbidity was identified using the Aetna Informatics Health Profile Database, used to identify Aetna members with chronic diseases or medical conditions. The identification algorithms are comprised of medical, pharmacy, and clinical laboratory data from physician claims and encounters, specialist claims, pharmacy, health care facilities, laboratories, and others (Hanchak, Hirsch, Murray, Schlackman, & McDermott, 1996). Specific comorbidities included gastritis or dyspepsia, hyperlipidemia, low back pain, chronic

TABLE 1 Summary comparison of key baseline sociodemographic and clinical features of tested individuals with full patient cohort, and with propensity-score matched cohort

		Matched groups				
	Baseline subject characteristics	**Matched control group**	**Tested group**	***P*-value†**	**Norm. diff.**	**Fixed match**
	N	2745	817			
Baseline characteristics	Age, mean (SD)	39.1(7.6)	40.7(13.3)	0.002	**0.115**	
	Gender (Female)	65.10%	65.10%	1.000	**0.007**	**Y**
	Median Household Income, median (Q1-Q3)	$84,029 ($64,704–$106,724)	$82,265 ($63,354 –$4104,544)	0.310	**0.163**	
	Geography: Region (Rural versus Urban)	20.30%	22.50%	0.273	**0.054**	
	Insurance type (Group, versus Individual)	23.50%	23.50%	1.000	**0**	**Y**
Clinical Characteristics	*Primary diagnosis (ICD9)*			1		**Y**
	Major Depressive Disorder	60.22%	60.22%			
	Anxiety Disorder*	24.97%	24.97%			
	Bipolar Disorder	14.81%	14.81%			
	Any anxiety diagnosis	53.49%	53.49%	1	**0**	**Y**
	Any substance use diagnosis	0.98%	0.98%	1	**0**	**Y**
	Treatment Trials, median (Q1-Q3)	2.0(1.0–4.0)	3.0(2.0–4.0)	<0.0001	**0.208**	
	Augment Therapy Trials, median (Q1-Q3)	0.0(0.0–1.0)	0.0(0.0–1.0)	<0.0001	**0.286**	
	All-cause Outpatient visits, mean (SD) [min-max]	2.64(2.99)[0.0–162.0]	2.54(4.9)[0.0–82.0]	0.380	**0.013**	
	All-cause ER visits, mean (SD) [min-max]	0.64(0.84)[0.0–32.0]	0.42(1.51)[0.0–28.0]	<0.0001	**0.171**	
	All-cause Inpatient visits, mean (SD) [min-max]	0.31(0.52)[0.0–14.0]	0.18(0.71)[0.0–6.0]	0.0001	**0.186**	
	Medical comorbidity					
	Hyperlipidemia	19.37%	21.18%	0.360	**0.042**	
	Low Back Pain	19.74%	20.20%	0.820	**0.012**	
	Hypertension	13.94%	15.79%	0.294	**0.046**	
	Migraine and Other Headaches	13.32%	13.46%	0.930	**0.004**	
	Diabetes Mellitus	5.92%	6.00%	0.940	**0.003**	
	Any mental health visit	53.49%	53.49%	1.000	**0**	**Y**
Baseline cost	6 months baseline paid claims			0.084	**0.052**	

thyroid disease, allergy, hypertension, migraine or other headache, or diabetes mellitus (Hanchak et al., 1996). In addition, baseline utilization was measured in terms of inpatient hospitalizations, outpatient visits, emergency room visits, and medication prescriptions. Psychiatric utilization was measured in terms of psychotropic medication prescriptions (Supporting Information Table 2) as well as emergency department visits and inpatient hospitalizations with a primary psychiatric diagnostic code. The 6-month lead-in and outcome periods were selected a priori to maximize follow-up while maintaining adequate sample size to allow application of propensity score matching.

2.3 | Propensity-score matching

One of the most common methods to adjust for baseline characteristics in observational studies is to match on the probability of receiving the treatment conditional on the characteristics. In settings with many baseline characteristics, propensity score matching is feasible when matching on all baseline characteristics directly is not feasible; comparing treated and control individuals with the same values for the propensity score removes the biases associated with all baseline characteristics. Here, we match exactly on a subset of the baseline characteristics (summarized below), and include others in the propensity

score. We then assess the balance by calculating normalized differences in baseline characteristics in the matched sample, calculated as the difference in means scaled by their standard deviation (Imbens & Rubin, 2015).

As advised by standard statistical texts, propensity score development proceeded iteratively considering different combinations of variables and matching conditions, while remaining blinded to outcome data (Imbens & Rubin, 2015). The aim was to estimate the probability of being in the study group (i.e., receiving assay-guided treatment) based upon baseline sociodemographic and clinical features, using logistic regression. Models included age, household income, coverage type, geography (urban vs. rural setting), baseline utilization, and individual baseline comorbidity. Models also included specific features of psychiatric treatment, including presence or absence of augmentation (i.e., addition of a non-antidepressant), number of pharmacologic trials, and days since last fill during the baseline period. For each candidate propensity score, normalized difference between experimental and control group (i.e., in means for continuous measures, or frequency for categorical ones) were compared for each variable. Per protocol, optimal propensity scores were also examined in terms of real prior-6-month all-in health care cost, although such costs were not released by the insurer for publication.

Following estimation of the optimal propensity score, each tested individual (case) was matched to up to four unique individuals from the control cohort. In particular, controls were required to match exactly on the following fixed matching criteria: gender; insurance type (employer or self-pay); presence or absence of an anxiety disorder; presence or absence of a substance use disorder; primary diagnosis of depression, bipolar disorder, or anxiety (Table 1 and Supporting Information Table 1); receipt of at least one visit in the prior 6 months for a mental health-related service, month of index visit, and log odds of propensity score ± 1. More precisely, first we ordered the control individuals in terms of a decreasing propensity score. Then for the first tested individual, we selected the control individuals closest in terms of the propensity score, who were matched exactly in terms of the fixed features already noted.

2.4 | Outcomes and analysis

Primary protocol-specified utilization outcomes during the 6-month follow-up period included count of emergency room visits, count of inpatient visits (overall, mood-disorder-related, and excluding mood-disorder-related), outpatient visits, and change in count of psychotropic drug prescriptions.

For each of these, we examined simple difference in average outcome (i.e., comparison in means) between matched case and control groups. To examine the robustness of these effects, we also applied negative binomial or zero-inflated negative binomial regression, as appropriate, adjusting for baseline sociodemographic and clinical features. For change in number of psychotropic medications, multiple linear regression was used. For ease of interpretability, all results are presented first in terms of simple two-group comparisons, then model-based (adjusted) results.

For each element of utilization, cost was estimated using the median 2012 cost from the Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project in order to account for variation by region, and maximize generalizability. (For outpatient visits, median cost was drawn from a 2008 Agency for Healthcare Research and Quality brief, with values adjusted for annual average inflation of 3.4% for 2008–12; Agency for Healthcare Research And Quality, 2008). All costs were summed to yield overall health care cost during the 6-month follow-up period. Test costs were excluded as the goal was to understand under what circumstances (i.e., testing cost) such testing will be cost effective.

In addition to simple two-group comparisons, cost analyses utilized generalized linear models with gamma distribution (log link) or two-part models with logistic regression followed by log-linear model with gamma distribution, as appropriate for the observed distribution. For two-group comparisons, *P*-values used nonparametric (Wilcoxon rank-based) test; in all cases, *P*-values were also significant for parametric (*t*-test) analysis.

All analyses utilized SAS version 9.4 and SAS Enterprise Guide version 6.1 (SAS Institute; Cary, NC).

3 | RESULTS

A total of 817 individuals out of 1,639 receiving AGT could be propensity-score matched to control individuals, based on the exact matching requirements already noted, yielding a control cohort of 2,745 individuals (Figure 1 and Table 1). Characteristics of the two matched groups at baseline, along with normalized mean differences, are described in Table 1. For comparison, Supporting Information Table 2 also contrasts tested individuals with the full control cohort ($n = 904{,}927$). As expected, the case and control groups were substantially more similar after matching, with modest remaining differences.

Matched individuals were first compared on measures of 6-month utilization (Table 2). In bivariate comparisons as well as model-based analyses, the AGT group had significantly fewer emergency room visits and inpatient hospitalizations than the controls. Among these hospitalizations, a more marked difference was observed for nonpsychiatric hospitalizations. (In a post hoc analysis, psychiatric emergency room visits were $\sim 2.5\times$ more common among untested individuals in model-based analysis; $P = 0.0002$). Outpatient visits were significantly less in the AGT group in model-based comparisons, but not in unadjusted bivariate tests. The two groups did not differ significantly in length of hospital stay or change in total number of psychotropic prescriptions.

Next, costs were imputed for individual categories of utilization as well as for overall 6-month utilization. These results are presented in Table 3, including model-based analyses (Table 3, right) and unadjusted bivariate comparisons (Table 3, left, and Figure 2). Overall costs were \$1,229 (SE 227) less in the AGT group in unadjusted comparisons, and \$1,948 (SE 611) less in fully adjusted models. (These costs do not include cost of the test, estimated to be \$750 for contracted health plans.) Post hoc diagnostics examining the fully adjusted model

TABLE 2 Measures of health care utilization among tested individuals and propensity-score matched controls in 6 months of follow-up

	Tested group		Matched control group								
	Mean	SE	Mean	SE	Unadjusted difference	SE	*t*-value	*P*-value	Model estimate	SE	% difference
Emergency Room Visits	0.19	0.02	0.33	0.02	0.13	0.03	4.10	<0.0001	0.14	0.02	40.4%
Inpatient hospitalization	0.07	0.01	0.17	0.01	0.10	0.02	4.78	<0.0001	0.17	0.04	57.9%
Inpatient hospitalization–non-mood disorder	0.05	0.01	0.14	0.01	0.09	0.01	5.40	<0.0001	0.14	0.02	65.5%
Inpatient hospitalization–mood disorder	0.03	0.01	0.03	0.00	0.01	0.01	2.04	0.04	0.01	0.00	22.8%
Outpatient Visits	1.29	0.09	1.47	0.06	0.19	0.10	3.64	0.0003	0.23	0.03	12.9%
Length of stay	8.11	1.14	9.14	0.67	1.04	1.21	0.20	0.84	1.16	0.78	11.4%
Change in number of psychotropic drug prescriptions	0.05	0.08	0.05	0.04	0.00	0.07	-1.09	0.28	0.00	0.03	8.0%

TABLE 3 Estimated health care costs among tested individuals and propensity-score matched controls in 6 months of follow-up

	Tested group		Matched control group								
	Mean	SE	Mean	SE	Mean difference	SE	*t*-value	*P*-value (*t*-test)	*P*-value (Wilcoxon)	Model estimate	SE
Emergency Room Visit Cost	$871	96	$1,521	94	$650	134	4.85	<0.0001	<0.0001	$599	$66
Inpatient Visit Cost	$443	86	$947	68	$505	103	4.6	<0.0001	<0.0001	$556	$23
Outpatient Visit Cost	$447	32	$474	18	$27	31	0.72	0.4724	0.2179	$36	$5
Pharmacy Cost	$378	78	$417	90	$39	123	0.32	0.7456	0.0135	$97	$54
Total Cost	$1,895	169	$3,123	158	$1,229	227	5.32	<0.0001	<0.0001	$1,948	$611

SE, standard error

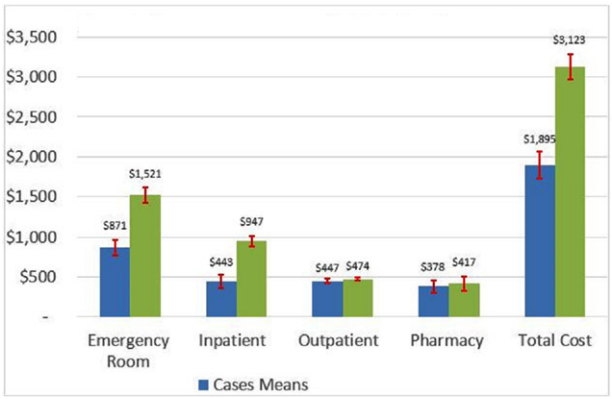

FIGURE 2 Illustration of mean cost among tested individuals and propensity-score matched controls in 6 months of follow-up

suggested outlier effects (see Supporting Information Figure 1); following Winsorization at the 99th percentile for cost, to minimize the impact of outliers, cost difference in the unadjusted model was $1047 (SE $161) and in the fully adjusted models $1498 (SE 82) less for the AGT group. The observed differences were largely the result of lesser 6-month emergency room costs ($650, SE 134) and inpatient costs ($505, SE 102) for the AGT group.

4 | DISCUSSION

In this case-control analysis of commercial insurance claims for 817 individuals with a mood or anxiety disorder who received assay-guided treatment, and 2,745 propensity-score matched individuals with the same diagnoses receiving standard-of-care treatment, testing was associated with significantly less utilization of emergency rooms and inpatient visits over the subsequent 6-month period. These differences translated to a significant difference in overall health care costs. Results were consistent in unadjusted bivariate comparisons of propensity-score matched individuals, and in fully adjusted regression models.

To our knowledge, these results represent one of the first demonstrations in a large patient cohort of clinically meaningful savings associated with a pharmacogenetic test; prior cost-effectiveness simulations had suggested potential benefit for single-locus pharmacodynamic assays in some circumstances (Perlis, Patrick, Smoller, & Wang, 2009; Olgiati, Bajo, Bigelli, De Ronchi, & Serretti, 2012). They are also consistent with a prior propensity-score matched cost-effectiveness analysis in 333 patients using the same assay investigated here, which found that outpatient costs in the 4 months following testing were 9.5% less than those among patients with TAU (Fagerness et al., 2014). The present study extends these results to

a large and generalizable national health care network, using more specific matching that accounts for baseline health care cost, and yielding more precise estimates of utilization and cost.

A recent analysis of health care costs associated with CYP450 substrate prescription in the state of Massachusetts suggested a potential mechanism that may contribute at least in part to this benefit (McCoy et al., 2017). That study found that CYP450 substrate medications are associated with $397/month increase in health care costs and (in two cohorts) a 10–13% increase in odds of 90-day hospital readmission (McCoy et al., 2017). Similarly, a small study among antidepressant-treated patients undergoing pharmacogenomic testing specifically indicated greater utilization (in terms of number of visits) among individuals with possible gene–drug interactions (Winner, Allen, Altar, & Spahic-Mihajlovic, 2013).

A key feature of the present study is generalizability: by utilizing broad inclusion criteria and robust propensity score based matching, these estimates should be widely applicable to US populations, i.e., should have great external validity relative to standard randomized trials (Steckler & McLeroy, 2008). In addition, the availability of comprehensive utilization data allows confidence that all costs are reflected in these analyses, minimizing the bias associated with other forms of data such as electronic health records where data may not be missing at random.

Still, we note multiple limitations that bear on interpretation of our results. First, only a double-blind, randomized trial allows truly unbiased estimates of efficacy. Here, the use of propensity score matching does diminish, but cannot guarantee elimination of, bias. In general, we would anticipate that noise introduced by baseline differences not captured in the propensity score would bias these results toward the null hypothesis. For example, if physicians who order the test are simply more aggressive or conscientious, we might expect a greater number of outpatient visits and pharmacy utilization in the tested group. Likewise, if testing reflects greater underlying illness severity, typically not well captured in health claims data, we would expect greater rather than lesser utilization among tested patients (Perlis et al., 2012). In particular, if selecting control individuals at any point in treatment (rather than at time of treatment change) biased us towards a more severely ill case population, we would expect to see greater cost among tested rather than untested patients. On the other hand, it is possible that clinicians order the test for less ill patients, while there is no evidence that this is the case based on cost prior to testing, if it were not addressed by propensity score matching this could falsely inflate the observed benefit.

Second, the present study does not allow direct measurement of clinical efficacy and quality of life using standardized instruments or scales, another potential source of confounding. Still, the nature of the intervention (a pharmacogenetic test, rather than a medication or device) and the improvement in measures that reflect adverse outcomes, such as emergency room visits, make it unlikely that the observed savings are obtained through diminished quality such as simply withholding treatment. If anything, decreasing intensity of care in a manner that reduces quality would be expected to increase utilization of emergency rooms and hospitalization. Although we controlled for geography, plan type, and clinical setting, provider-level data were not available and these variables might explain additional variance in outcomes. However, testing was not restricted to any particular contract type or group of clinicians. If ordering clinicians are more skilled or 'better' in some way, it is hard to envision why this would be the case only after ordering a test. Finally, the nature of clinical data necessitates numerous assumptions about variable distribution for statistical analysis. Here, the consistency of results between simple bivariate tests and regression models in sensitivity analysis suggests that these findings are robust to common assumptions.

Finally, we cannot distinguish test-specific effects from expectancy or placebo-like effects: tested individuals would be likely to anticipate benefit of test-guided treatment. Prior work suggested that pharmacogenetic testing with this assay increased adherence, but the present data did not allow us to examine the extent to which benefits may be mediated by improved adherence (McCoy et al., 2017). On the other hand, the reduction of 'hard' utilization outcomes, rather than, for example, clinician-assessed severity, suggests that placebo effects or improvement in adherence would need to be substantial. To date, there are no true double-blind trials of pharmacogenetic tests that blind clinician, patient, and evaluator; such trials pose substantial logistic and ethical constraints that merit further consideration and methodologic innovation.

5 | CONCLUSION

Taken together, these results add to a growing body of evidence that pharmacogenetic testing as a guide to psychiatric prescribing is associated with a longer-term reduction in health care utilization and costs. Additional study will be useful in further defining the optimal application of such testing in clinical practice, and the impact of testing on quality as well as quantity of care.

ACKNOWLEDGMENTS

R.H.P. has served on advisory boards or consulted to Genomind, Healthrageous, Perfect Health, Pfizer, Psy Therapeutics, and RIDVentures. R.M., A.M.E., and A.T. are employees of Healthagen, an Aetna subsidiary. G.W.I. has consulted for Genomind, Eli Lilly and Company, and Amazon. The study was supported by Genomind.

CONFLICT OF INTEREST

Dr. Imbens reports personal fees from Genomind, for experimental design related to the submitted work, and personal fees from Eli Lilly, outside the submitted work. Dr. Perlis reports grants from the National Human Genome Research Institute and grants from the National Institute of Mental Health; personal fees from Genomind for experimental design related to the submitted work, personal fees from Healthrageous, personal fees from Perfect Health, personal fees from Psy Therapeutics, and personal fees from RID Ventures, outside the submitted work. All other authors reported no disclosures.

ORCID

Roy H. Perlis MD, MSc http://orcid.org/0000-0002-5862-6757

REFERENCES

Agency for Healthcare Research And Quality (2008). "National Healthcare Quality Report", Retrieved from https://archive.ahrq.gov/research/findings/nhqrdr/nhqr08/index.html.

Brennan, F. X., Gardner, K. R., Lombard, J., Perlis, R. H., Fava, M., Harris, H. W., & Scott, R. (2015). A naturalistic study of the effectiveness of pharmacogenetic testing to guide treatment in psychiatric patients with mood and anxiety disorders. *The Primary Care Companion for CNS Disorders*, *17*, https://doi.org/10.4088/PCC.14m01717

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*, e146–156.

Federal Drug Administration (2016). "Table of pharmacogenomic biomarkers in drug labeling", Retrieved from https://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm.

Hanchak, N. A., Murray, J. F., Hirsch, A., McDermott, P. D., & Schlackman, N. (1996). USQA Health Profile Database as a tool for health plan quality improvement. *Management Care Quarterly*, *4*, 58–69.

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference for statistics, social, and biomedical sciences*. New York, NY: Cambridge University Press.

Institute of Medicine of the National Academies (2010). Policy issues in the development of personalized medicine in oncology. Workshop summary. Washington, D.C.: National Academies Press.

McCoy, T., Castro, V., Cagan, A., Roberson, A. M., & Perlis, R. H. (2017). Prevalence and implications of cytochrome P450 substrates in Massachusetts hospital discharges. *The Pharmacogenomics Journal*, *17*, 382–385.

Office of the Press Secretary (2015). *FACT SHEET: President Obama's precision medicine initiative. Statements and releases*, Washington, D.C.: The White House.

Olgiati, P., Bajo, E., Bigelli, M., De Ronchi, D., & Serretti, A. (2012). Should pharmacogenetics be incorporated in major depression treatment? Economic evaluation in high-and middle-income European countries. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *36*, 147–154.

Perlis, R., Iosifescu, D., Castro, V. M., Murphy, S. N., Gainer, V. S., Minnier, J., …, Smoller, J. W. (2012). Using electronic medical records to enable large-scale studies in psychiatry: Treatment resistant depression as a model. *Psychological Medicine*, *42*, 41–50.

Perlis, R. H., Patrick, A., Smoller, J. W., & Wang, P. S. (2009). When is pharmacogenetic testing for antidepressant response ready for the clinic? A cost-effectiveness analysis based on data from the STAR* D study. *Neuropsychopharmacology*, *34*, 2227–2236.

Steckler, A., & McLeroy, K. R. (2008). The Importance of External Validity. *American Journal of Public Health*, *98*, 9–10.

United States Census Bureau "American community survey", Retrieved from https://www.census.gov/programs-surveys/acs/.

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

Zhou, S. F., Liu, J. P., & Chowbay, B. (2009). Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metabolism Reviews*, *41*, 89–295.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Perlis RH, Mehta R, Edwards AM, Tiwari A, Imbens GW. Pharmacogenetic testing among patients with mood and anxiety disorders is associated with decreased utilization and cost: A propensity-score matched study. *Depress Anxiety*. 2018;1–7. https://doi.org/10.1002/da.22742