Research proposal

The social cost of suboptimal medication use and the value of pharmacogenomic information: evidence from Geisinger Health System

Frank R. Lichtenberg, Rebecca Pulk, Marc S. Williams, and Eric Wright

Abstract

Many drugs don't work the same way for everyone, but it is often difficult to predict who will benefit from a medication, who will not respond at all, and who will experience adverse drug reactions, which are a significant cause of hospitalizations and deaths in the United States. This study will use a unique database, consisting of a 10-year history of clinical and claims data on tens of thousands of patients linked to their genetic information obtained by whole exome sequencing, to estimate the social cost of suboptimal medication use and the value of pharmacogenomic information. The study will be conducted by an interdisciplinary team comprised of a health economist, a physician/medical geneticist, and two doctors of pharmacy.

Background

Pharmacogenomics has the potential to optimize clinical outcomes by streamlining the selection of medications to maximize benefit and minimize risk for an individual patient. This subsection of precision medicine has the capacity to significantly decrease medical expenditure by maximizing the clinical effectiveness of commonly prescribed medications while minimizing both medication nonadherence and adverse drug events. The potential clinical utility of pharmacogenomics can be observed by the steady increase in the number of practice guidelines published to direct the use of medications based on genetic markers [1, 2]. In spite of the touted clinical benefits of pharmacogenomics, the adoption of this technology into routine practice has been limited by the dearth of real-world evidence on the clinical and economic benefits of pharmacogenomics-guided prescribing.

Limited data linking patient-specific genomic and clinical information are available to inform clinical practice and coverage determinations [3]. Only a small number of health systems have integrated routine application of pharmacogenomics into their clinical workflows or pharmacogenomic test results into their medical records. Even fewer data repositories tie economic information to genomic and clinically-linked patient outcomes. The majority of the analyses concerning the value of pharmacogenomics-based prescribing have relied on modeling methods. These methodologies are limited by their assumptions, choice of comparator, and time frame, as well as their societal perspective. Assessing the value of pharmacogenomics-guided prescribing is further confounded in that these tests have historically been covered under prescription drug benefits, limiting the extent to which testing can be linked to total cost of care. These limitations have discouraged health systems from investing in implementation of pharmacogenomics.

In order for pharmacogenomics-based prescribing to become commonplace, the clinical and economic utility of this technology must be established.

Project Aims

The analysis we propose to perform has two main objectives:

1) To describe the clinical impact of pharmacogenomics-guided prescribing for three or more specific drug/drug class-gene pairs within a real-world healthcare environment
2) To quantify the economic impact of pharmacogenomics-guided prescribing based on these findings

Methods

Study Design. We will conduct a retrospective analysis utilizing data from the Geisinger Electronic Health Record (EHR) and Geisinger Health Plan claims database (which includes prescription drug claims) over a 10-year period. All patients who have consented to genomic analysis and have received a medication of interest since January 1, 2006 will be considered for inclusion.

Setting/data sources. As one of the nation's largest health care organizations, Geisinger serves more than 3 million residents throughout 45 counties in central, south-central and northeast Pennsylvania, as well as in southern New Jersey with the addition of AtlantiCare, a National Malcolm Baldridge Award recipient. The Clinic's patient population is very stable: census data indicate that with the exception of two counties, the out-migration rate is less than 1% per year. The physician-led system is comprised of approximately 30,000 employees, including nearly 1,600 employed physicians, 11 hospital campuses, 83 primary and specialty clinic sites including 41 community-based primary care clinics, two research centers and a 510,000-member health plan.

Geisinger Health System is an integrated health care organization widely recognized for its development of high reliability care delivery models, such as ProvenHealth Navigator® and ProvenCare®, and innovative use of the electronic health record. Geisinger has improved health outcomes (e.g. reduced hospitalizations and amputations) through an integrated network of providers, oversight by embedded case nurse managers, and systematic initiatives designed to optimize patient-oriented outcomes [4, 5, 6, 7]. As a system they have embraced research as a means to improve patient care. Geisinger's EHR, Epic, contains information on more than 3 million unique patients and has been in use since 1999 in all outpatient clinics, and since 2007 in hospitals. To date, our EHR database contains information on more than 4 million patients with over 600,000 unique patients having encounters in the health system each year. Information from the electronic health record and medical/pharmacy claims are stored in data warehouses accessible to both the clinical and research enterprises.

Geisinger has made significant investments in MyCode®, a community health initiative that links longitudinally collected biobank samples to initiatives aimed at leveraging genomic and phenotypic information. The MyCode® project currently has over 100,000 participants who have

consented to participate. Consent allows broad research use and recontact and includes whole exome sequencing (WES) of their genome and is facilitated in a partnership (DiscoverEHR™) with Regeneron Pharmaceuticals Inc. DiscoverEHR™ has run and quality-controlled over 50,000 patients for WES, with approximately 1000 additional patients consenting each week. The goal is to enroll 250,000 patients in the next five years. Projects focusing on discovery and optimization of outcomes are underway to capitalize on the growing reservoir of information.

WES is a genomic sequencing technology that focuses on the information contained in the 1% of human DNA that codes for proteins. This technology, while very powerful for detecting single nucleotide variants and small insertions and deletions (indels), is limited in its ability to detect larger copy number variants and areas of the exome that have repetitive motifs. It does not capture variation for the 99% of the genome outside of these target regions. An exome done as part of a research project or for a diagnostic indication has the potential to yield pharmacogenomic (PGx) data "for free". To date it has been difficult to directly extract actionable PGx information from WES data, thus literature assessing the cost effectiveness of WES in directing pharmacotherapy is unavailable. Recent advances in database mining techniques will enable extraction of PGx markers from Geisinger's WES data.

Baseline Information. We aim to study the following drug-gene pairs: statins/*SLCO1B1*, clopidogrel/*CYP2C19*, and antidepressants/*CYP2C19*. Based upon data collected from January 1, 2006 to April 15, 2016, we have identified significant populations of patients on these the medications of interests with available WES genomic data, approximately forty percent of which also carry GHP insurance. **Table 1** below provides a detailed count of our internal inquiry, including data for additional drug-gene pairs that may allow for further analysis.

Inclusion criteria. For this project we will utilize clinical and genetic data from Geisinger's MyCode® Community Health Initiative, as well as Geisinger Health Plan (GHP) claims data to assess adherence to our medications of interest and to measure utilization for covered members that occurs outside the Geisinger system. These data will allow us to assess medication persistence, adverse drug reactions and total cost of care.

Aim 1 analysis. The breadth and depth of the Geisinger dataset will allow us to explore the clinical impact of pharmacogenomic guided prescribing in a real world setting. We will limit our analysis to drug - germline gene pairings with available prescribing guidelines from one of the following organizations: Clinical Pharmacogenomics Implementation Consortium (CPIC), Dutch Pharmacogenomics Working Group (DPWG) and the Canadian Pharmacogenomics Network for Drug Safety (CPNDS). We will be utilizing the genetic information available from the Geisinger MyCode® Community Health Initiative for members who have consented to WES linked to electronic health record data. This rich source of data will allow us to examine the effects of multiple drug-gene pairs simultaneously.[1]

[1] Since our source of genomic information (whole exome sequencing) only provides information on an individual patient's native genetic sequence, we will not examine the cost-effectiveness of screening for tumor markers. We will also exclude the gene *CYP2D6* and all HLA markers associated with adverse events due to the technical difficulties of extracting actionable information from WES.

We will compare the clinical outcomes of patients prescribed medications that would be judged 'optimal' on the basis of their pharmacogenomics markers against those whose genetic markers indicate an alternative medication or dose would have been more optimal. Specific clinical outcomes of interest for each drug-gene pair studied can be found in Table 1. General clinical measures that will be measured for all subpopulations will include the rate of hospitalization, rate of ED visits and outpatient visit rate. We will also report the number of relevant pharmacogenomics findings (for specified drug/gene pairs) for each subject included in the analysis.

Aim 2 Analysis. The Geisinger dataset will further allow us to quantify the economic impact of pharmacogenomics-guided prescribing in a real-world setting. Following our determination of the clinical outcomes, we will utilize claims data to evaluate the financial implications of utilizing a medication against the available pharmacogenomic guided prescribing guidelines. We will specifically quantify the average cost differential seen between patients receiving optimal versus suboptimal therapy based on their pharmacogenomic markers. For patients not covered under GHP, we will be able to estimate cost savings from event rates and average costs of care.

Outcomes. Outcomes data will be extracted from both the electronic health record and the GHP claims data repository. As shown in Table 2, clinical outcomes of interest will include medication persistence, frequency of adverse events and health care utilization. Economic outcomes of interest include the average total cost of care for each subpopulation identified. Secondary economic outcomes of interest include the cost effectiveness of pharmacogenomic testing.

Statistical Analysis. We will estimate models of the following form:

$$OUTCOME_i = \beta\ SUBOPTIMAL_i + \gamma\ Z_i + \varepsilon_i$$

where

$OUTCOME_i$ = an outcome measure (e.g. medication persistence, frequency of adverse events, or health care utilization) for patient i

$SUBOPTIMAL_i$ = 1 if one or more medications administered to patient i was suboptimal for the patient, based on subsequently obtained PGx data
= 0 otherwise

Z_i = an extensive set of other patient characteristics, e.g. age, sex, and comorbidities

ε_i = a disturbance

Patient Protection. This study will receive approval from the Geisinger IRB before starting, and all patient data will be extracted from the databases by Geisinger's Biostatistics Core programmers who have been trained in data security and patient confidentiality. Identifiable data will be shared only with the IRB-approved study team members and those with approved data-use agreements. All patient information will be kept in password-protected folders on secure file servers.

Timeline. The timeline for this project is 12 months from initiation to completion.

Deliverables. Findings will be presented at national meetings and in peer-reviewed journal(s).

Research Team

Frank R. Lichtenberg, PhD, is Courtney C. Brown Professor of Business at the Columbia University Graduate School of Business; a Research Associate of the National Bureau of Economic Research; and a member of the CESifo Research Network. He received a BA with Honors in History from the University of Chicago and an MA and PhD in Economics from the University of Pennsylvania.

Rebecca Pulk, MS, PharmD (expected June 2016) received a BS in Chemistry from Pepperdine University and an MS in Pharmaceutical Chemistry from the University of Florida. She is a Doctor of Pharmacy candidate at the University of Minnesota. She will be joining Geisinger's Center for Health Research as a postdoctoral fellow in July of 2016 following graduation and will be focused on the implementation and outcomes of pharmacogenomics-guided prescribing.

Marc S. Williams, MD, FAAP, FACMG is the director of the Geisinger Health System Genomic Medicine Institute. His primary research interest has been to study the impact of implementing genetics, genomics and family history in routine clinical care. He has published several papers related to economic evaluation of pharmacogenomics, and co-authored the book *Economic Evaluation of Genomic Medicine*.

Eric Wright, PharmD, MPH is a director and senior investigator at Geisinger's Center for Pharmacy Innovation and Outcomes. He received his Doctor of Pharmacy from the University of Pittsburgh and Masters of Public Health from Johns Hopkins. His primary research focus has been to evaluate and optimize medication use through improved use of new information, leveraging of health information technology, and progressive integration of pharmacy services.

References

[1] Dosing Guidelines. Pharmacogenomics Knowledgebase (PharmGKB), Accessed April 25, 2016: https://www.pharmgkb.org/view/dosing-guidelines.do

[2] Publications Clinical Practice Guidelines / Recommendations: Canadian Pharmacogenomics Network for Drug Safety, Accessed April 25, 2016 http://cpnds.ubc.ca/about/projects

[3] Snyder SR, Mitropoulou C, Patrinos GP, Williams MS. Economic evaluation of pharmacogenomics: a value-based approach to pragmatic decision making in the face of complexity. Public Health Genomics. 2014;17(5-6):256-64. doi: 10.1159/000366177. Epub 2014 Sep 26.

[4] Gilfillan RJ, Tomcavage J, Rosenthal MB, et al. Value and the medical home: Effects of transformed primary care. Am J Manag Care. 2010;16(8):607-614.

[5] Steele GD, Haynes JA, Davis DE, et al. How Geisinger's advanced medical home model argues the case for rapid-cycle innovation. Health Affairs. 2010;29(11):2047-2053. doi: 10.1377/hlthaff.2010.0840.

[6] Maeng DD, Graf TR, Davis DE, Tomcavage J, Bloom FJ. Can a patient-centered medical home lead to better patient outcomes? the quality implications of Geisinger's ProvenHealth navigator. American Journal of Medical Quality. May/June 2012;27(3):210-216. doi: 10.1177/1062860611417421.

[7] Maeng DD, Khan N, Tomcavage J, Graf TR, Davis DE, Steele GD. Reduced acute inpatient care was largest savings component of Geisinger health system's patient-centered medical home. Health Affairs. 2015;34(4):636-644. doi: 10.1377/hlthaff.2014.0855.

Table 1. Cross-sectional counts of patients within Geisinger prescribed listed drugs along with counts of patients with MyCode® information and GHP insurance and measures of interest

Drug	N patients prescribed in past 10 years with MyCode® samples	% with GHP	Variant needed	Population penetrance (% with pharmaco-pathogenic variant)	Potential outcomes of interest
Simvastatin, Atorvastatin or Rosuvastatin	33344	42.77%	*SLCO1B1* C allele rs4149056	Caucasians: 10-20% AAs: 3%	Myalgia, myopathy and rhabdomyolysis Statin persistence CV outcomes / events
Clopidogrel	9049	39.44%	*CYP2C19* LOF markers: *2 rs4244285 *3 rs4986893 *4 rs28399504 *5 rs56337013 *6 Haplotype (3 SNPs) *7 rs72558186 *8 rs41291556	Caucasians: PMs:2% IMs: 25% AA: PMs: 5% IMs:30%	Cardiovascular outcomes / events Cost of care
Escitalopram Citalopram Sertraline	11135	42.88%	*CYP2C19* LOF markers: *2 rs4244285 *3 rs4986893 *4 rs28399504 *5 rs56337013 *6 Haplotype (3 SNPs) *7 rs72558186 *8 rs41291556 GOF marker: *17 rs12248560	Caucasians: PMs:2% IMs: 25% AA: PMs: 5% IMs:30% General: URMs: 3.8%	Persistence/ adherence Symptoms and/ or remission Health care utilization (visits, labs, hospitalizations)

Omeprazole, Esomeprazole, Pantoprazole, or Rabeprazole	19443	43.00%	*CYP2C19* LOF markers: *2 rs4244285 *3 rs4986893 *4 rs28399504 *5 rs56337013 *6 Haplotype (3 SNPs) *7 rs72558186 *8 rs41291556 GOF marker: *17 rs12248560	Caucasians: PMs:2% IMs: 25% AA: PMs: 5% IMs:30% General: URMs: 3.8%	GOF H pylori treatment failure LOF QTc prolongation INR increase
Voriconazole	82	41.57%	*CYP2C19* LOF markers: *2 rs4244285 *3 rs4986893 *4 rs28399504 *5 rs56337013 *6 Haplotype (3 SNPs) *7 rs72558186 *8 rs41291556 GOF marker: *17 rs12248560	Caucasians: PMs:2% IMs: 25% AA: PMs: 5% IMs:30% General: URMs: 3.8%	LOF: Drug toxicity markers: rash, transaminitis, jaundice, fever, serum creatinine increase GOF: treatment failure
Mercaptopurine Thioguanine Azathioprine	690	41.51%	*TPMT* *2 *3A *3B *3C *4	Low Function 10% of patients Absent: ~1 in 400 patients	Neutropenia Or neutropenic fever Clinical outcomes Health care costs

Capecitabine 5FU	249	35.38%	*DPYD* *2 *2A *13 rs67376798	Low function ~ 4% of patients No function ~0.5% of patients	Mucositis, diarrhea, nausea, altered mental status, acute renal failure, and rashes.
Tacrolimus	375	32.22%	*CYP3A5* *1 (many) *6 rs10264272 *7 rs76293380	Caucasians: *1: 8% *6: 0.1% AA: *1: 56% *6: 17% *7: 8%	Labile drug level troughs Episodes of acute rejections

LOF = Loss of function (decreased metabolic activity)
GOF = Gain of function (increased metabolic activity)
SNP = Single nucleotide polymorphism
AA = African American
PM = Poor metabolizer
IM = Intermediate metabolizer
URM = Ultra-rapid metabolizer

Table 2. Measurement of clinical outcomes

Clinical outcome	Measurement of outcome
Myalgia, Myopathy and rhabdomyolysis	Laboratory measures of CPK, a discrete marker for muscle damage, will be extracted from the EMR. EMR notes and documents will be screened for notes mentioning adverse drug reactions (non-discrete). Claims will be screened for ICD codes associated with myopathy or rhabdomyolysis.
Statin Persistence	Prescription claims data will be used to assess medication adherence through medication possession ratios.
Cardiovascular Outcomes / Events	EMR and claims data can be screened to identify hospitalizations due to major cardiovascular adverse events including MI and stroke by ICD code. EMR data can be utilized to measure mortality.
Cost of care	Direct measurement of cost of care for patients covered by GHP will be derived from claims data.
Persistence / adherence	Prescription claims data will be used to assess medication adherence through medication possession ratios.
Depressive symptoms and/or remission	PHQ9 scores, a discrete patient reported measure of depressive symptoms, will be extracted from the medical record.
Health care utilization (visits, labs, hospitalizations)	Estimation of health care utilization will be made through query of the EMR and GHP claims data.

CPK = creatine phosphokinase
ICD = international classification of disease
PHQ9 = patient health questionnaire