2018

A Multifactorial Cytochrome P450 2D6 Genotype-Phenotype Prediction Model to Improve Precision of Clinical Pharmacogenomic Tests

Rachel Dalton
University of Montana - Missoula

A MULTIFACTORIAL CYTOCHROME P450 2D6 GENOTYPE-PHENOTYPE PREDICTION MODEL TO IMPROVE PRECISION OF CLINICAL PHARMACOGENOMIC TESTS

By

RACHEL NOELLE DALTON

PharmD, University of Montana, Missoula, Montana 2017
BA Biological Sciences, University of Chicago, Chicago, Illinois, 2010

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Science
in Pharmaceutical Sciences

The University of Montana
Missoula, MT

Official Graduation Date: Fall 2017

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

Erica L. Woodahl, PhD, Chair
Department of Biomedical and Pharmaceutical Sciences

Richard Bridges, PhD
Department of Biomedical and Pharmaceutical Sciences

Keith Parker, PhD
Department of Biomedical and Pharmaceutical Sciences

Nigel Priestley, PhD
Department of Chemistry and Biochemistry

Dalton, Rachel, M.S., Fall 2017 Pharmaceutical Sciences

A multifactorial cytochrome P450 2D6 genotype-phenotype prediction model to improve precision of clinical pharmacogenomic tests

Chairperson: Erica Woodahl

BACKGROUND: *CYP2D6* is difficult to accurately genotype due to a large number of single nucleotide variants (SNVs), indels, and structural variation such as deletions, duplications, and *CYP2D6/CYP2D7* hybrid genes. *CYP2D6* targeted genotyping panels are of limited utility; clinically relevant variants that are not genotyped will be missed. Sequencing solves this problem but requires additional tools to address structural variation. The goal of our study was to determine the predictive power of Stargazer, a novel allele-calling program, which combines SNV/indel calls with structural variation identification.
METHODS: In a panel of 309 human livers, *CYP2D6* diplotypes and activity scores were initially assigned manually using PGRNSeq SNV/indel data and then reassigned after inclusion of Stargazer-derived structural variation data. We determined CYP2D6 activity in human liver microsomes with metoprolol and dextromethorphan as probe substrates. Then, we used linear regression to assess the relationship between activity and activity scores assigned using SNV/indel data alone versus SNV/indel + structural variation data.
RESULTS: Without incorporating structural variation data, diplotypes were incorrectly assigned for 67 samples (22%); activity scores were incorrect for 26 samples (8.4%). Structural variants included 23 deletions, 47 duplications, and 39 hybrids. When diplotypes were assigned based on SNV/indel data alone, activity score explained 31% of the variation in CYP2D6 activity with metoprolol ($R^2 = 0.31$, $p < 0.001$) and 36% with dextromethorphan ($R^2 = 0.36$, $p < 0.001$). When reassigned with SNV/indel plus structural variation data, this increased to 36% for metoprolol ($R^2 = 0.36$, $p < 0.001$) and 41% for dextromethorphan ($R^2 = 0.41$, $p < 0.001$).
CONCLUSION: The accuracy of CYP2D6 phenotype prediction can be improved by using a next-generation sequencing approach coupled with a tool such as Stargazer to detect common and rare SNVs and indels as well as structural variation in *CYP2D6.*

Table of Contents

1 Introduction 6
1.1 Pharmacogenomics 6
1.1.1 Translation from bench to bedside: an unmet need 6
1.1.2 Drug-metabolizing enzymes and transporters 9
1.1.2.1 Phase I Drug Metabolism 10
1.1.2.2 Phase II Drug Metabolism 11
1.1.2.3 Drug Transport 11
1.1.3 Applications 12
1.1.3.1 FDA-required pharmacogenetic tests 13
1.1.3.2 Clinical Pharmacogenomics Implementation Consortium guidelines 14
1.1.3.3 Preemptive pharmacogenomics programs 15
1.2 Cytochromes P450 17
1.2.1 CYPs responsible for drug metabolism 17
1.2.2 Biochemistry 20
1.2.3 Principles of *in vitro* P450 reaction phenotyping 22
1.2.3.1 In vitro models of drug metabolism 23
1.2.3.2 Human liver microsome incubations 25
1.2.4 Allele nomenclature 27
1.2.5 Translating CYP diplotype to phenotype 28
1.3 CYP2D6 31
1.3.1 Clinical relevance 31
1.3.1.1 Codeine 32
1.3.1.2 Tricyclic antidepressants 33
1.3.1.3 Selective serotonin reuptake inhibitors 34
1.3.1.4 Emerging areas 35
1.3.2 CYP2D6 gene locus 36
1.3.3 CYP2D6 nomenclature 36
1.3.4 Activity score 37
1.3.5 Relevant *CYP2D6* alleles 38
1.3.6 Other sources of CYP2D6 variation 42
1.3.6.1 Ethnicity 42
1.3.6.2 Age and sex 43
1.3.6.3 Drug-drug interactions 43
1.3.6.4 Liver disease 43
1.3.6.5 Cytochrome P450 Oxidoreductase 43
1.3.6.6 Aldo-keto reductase 1D1 44
1.3.7 *CYP2D6* genotyping methods 44
1.4 Specific aims 46

2 A multifactorial cytochrome P450 2D6 genotype-phenotype prediction model to improve precision of clinical pharmacogenomic tests 47
2.1 Introduction 47
2.2 Materials and methods 49
2.2.1 Chemicals and reagents 49
2.2.2 Human liver samples 49
2.2.2.1 Human liver microsomes, total protein quantitation, cytochrome P450 oxidoreductase protein quantitation, genomic DNA extraction 49
2.2.2.2 CYP2D6 and POR protein quantitation and CYP2D6, POR, and AKR1D1 mRNA quantitation 49
2.2.2.3 DNA isolation, CYP2D6 sequencing, and CYP2D6 allele and activity score assignment 50
2.2.2.4 Measurement of CYP2D6 metabolite formation rate in HLMs 50
2.2.2.5 Metabolite and parent drug quantitation 51
2.2.2.6 Statistical analysis 52
2.3 Results 52
2.3.1 Liver donor demographics 52
2.3.2 Effects of CYP2D6 mRNA and protein content on activity 53
2.3.3 CYP2D6 allele and diplotype frequencies in the liver bank samples 53
2.3.4 Structural variation data is critical for correct diplotype assignment 54
2.3.5 Liver collection site, POR protein content, and *AKR1D1* mRNA content contribute to variability in CYP2D6 activity 56
2.3.6 Multivariate regression analysis of CYP2D6 activity 56
2.4 Discussion 57
2.5 Tables 63
2.5.1 Table 1. Liver donor demographics (n = 309) 63

2.5.2 Table 2. *CYP2D6* alleles (n = 618) identified in liver bank samples 64
2.5.3 Table 3. *CYP2D6* diplotypes (n = 309) identified in liver bank samples 65
2.5.4 Table 4. Multiple linear regression: comparison of association between CYP2D6 activity and activity score when activity score is assigned using SNV/indel data only and SNV/indel plus structural variation data .. 66
2.5.5 Table 5. Multiple linear regression: association between CYP2D6 activity and activity score, POR protein content, and *AKR1D1* mRNA content 67
2.6 Figures 68
2.6.1 Figure 1. Comparison of diplotype and activity score changes assigned with SNV/indel data alone and with inclusion of structural variation data with Stargazer 68
2.6.2 Figure 2. Association between CYP2D6 metabolite formation rate and CYP2D6 activity score. 69
2.6.3 Figure 3. Association between CYP2D6 metabolite formation rate and *CYP2D6* mRNA and protein content. 70
2.6.4 Figure 4. Association between CYP2D6 metabolite formation rate and POR protein and *AKR1D1* mRNA content 71

3 Discussion 72
3.1 Project summary 72
3.2 Emerging roles for pharmacists 75
3.2.1 Physician knowledge gaps 75
3.2.2 Pharmacist qualifications 76
3.2.3 Pharmacist roles in existing pharmacogenomics programs 78
3.2.4 Expansion beyond academic medical centers 79

4 Appendix: optimization of human liver microsome incubations 81
4.1 Selection of probe drugs and drug concentrations 81
4.2 Selection of HLM protein concentrations and incubation times 82

5 References 85

1 Introduction

1.1 Pharmacogenomics

Pharmacogenomics is the study of how an individual's genetics affect their response to drugs, in terms of efficacy and/or toxicity.[1] Clinically, pharmacogenomics is used to optimize drug therapy by guiding the selection of a drug and dose that will maximize efficacy and minimize adverse effects. Pharmacogenetic tests may be used to assess genetic variation in drug-metabolizing enzymes or transporters, components of the immune system that may contribute to idiosyncratic adverse drug reactions, or genes involved in a drug's mechanism of action. [2]

1.1.1 Translation from bench to bedside: an unmet need

The publication of Arno Motulsky's 1957 paper *Drug reactions, enzymes, and biochemical genetics* is generally regarded as the beginnings of pharmacogenomics as a discipline.[3] In this paper, he discussed "how hereditary gene-controlled enzymatic factors determine why, with identical exposure, certain individuals become 'sick,' whereas others are not affected." In 1959, two years after Motulsky's publication, Friederich Vogel coined the term "pharmacogenetics." The term "pharmacogenomics" has since been introduced and reflects advances in knowledge and technology that allows investigation of the entire genome. [1,3]

Despite a long history, the field of pharmacogenomics is only starting to move into the clinical setting. Both the United States Food and Drug Administration (FDA) and the National Institutes of Health (NIH) have recognized the potential behind precision medicine and have launched initiatives to support the necessary scientific and regulatory

structure.[4] One mechanism used by the FDA is inclusion of pharmacogenetic data on drug labels. Specifically, this includes information that "has important implications for safe and effective use and the consequences of the genetic differences result in recommendations for restricted use, dosage adjustments, contraindications, or warning."[5] Simply including this information on the label, however, may not be enough to encourage clinicians to use pharmacogenetic testing to guide clinical decision-making. In 2014, Wang et al. reviewed the FDA 199 drug labels that included pharmacogenetic data. Of these, only 43 (36.1%) provided convincing evidence of clinical validity. Additionally, only 61 (51.3%) of the labels made any recommendation about how to use the results of a pharmacogenetic test.[6] This study highlights the prerequisite need to provide clear, evidence-based recommendations to clinicians if they are to use pharmacogenetic tests in practice.

Studies investigating clinician's attitudes towards the utility of pharmacogenomics in clinical practice further demonstrate the need to refine the way pharmacogenomics data is presented. A 2015 survey of 90 Mayo Clinic primary care physicians found that 52% of the clinicians did not expect or know whether they would use pharmacogenomics information for future prescribing. Furthermore, 53% reported that the pharmacogenomics Clinical Decision Support alerts integrated into the Mayo Clinic's electronic health records software were, "confusing, irritating, frustrating, or difficult to use." Of those who received a Clinical Decision Support alert, only 30% altered their original prescription.[7] Even at a site such as the Mayo Clinic that has implemented a dedicated pharmacogenomics program, clinicians are often still not getting the information they need delivered in the right way.

Initial results from the University of Chicago's 1200 Patients Project suggest that a more streamlined process may improve acceptance of pharmacogenomics recommendations among physicians. The 1200 Patients Project uses the Genomics Prescribing System (GPS) to provide alerts to physicians. When there is a pharmacogenomics alert for a specific patient, the prescriber will see a graphic of a stoplight with a red, yellow, or green light, indicating the severity of the results. Upon clicking on the stoplight, a summary that can be read in 30 seconds or less describes the drug/gene interaction. Initial results published after 812 patients had been enrolled indicated that physicians clicked on the red light alert 100% of the time, the yellow light alert 72% of the time, and the green light alert 20% of the time. When surveyed after an alert prompted a medication change, 10% agreed strongly with the GPS recommendation, 76% agreed somewhat, 3% disagreed somewhat, and 10% disagreed strongly. Furthermore, 93% of respondents said that they were "very likely" to enroll more of their patients in the study.[8]

To address issues involved with efficient delivery of usable information, groups composed of clinicians, academics, and industry representatives, such as the Clinical Pharmacogenomics Implementation Consortium (CPIC), have been formed to with the goal of developing clinical practice guidelines for pharmacogenomics test interpretation. Similar groups exist internationally such as the European Pharmacogenetics Implementation Consortium (Eu-PIC) and the Dutch Pharmacogenetics Working Group (DPWG). The primary activity of the CPIC is to publish "freely available, peer-reviewed, updatable, and detailed gene/drug clinical practice guidelines" that "enable the translation of genetic laboratory test results into actionable prescribing decisions for specific drugs."[9]

As of January 2018, the CPIC has published guidelines for 36 different drug/gene combinations. Some of the guidelines are for a single drug/gene pair, such as codeine and *CYP2D6*, while others are for a drug and a combination of genes, such as warfarin and *CYP2C9* and *VKORC1*.[10] Each guideline conforms to a standardized format and system to grade levels of evidence and to score the strength of each prescribing recommendation. The guidelines contain a literature review, information on the gene of interest, a summary of available genetic tests, guidance on how to interpret genetic tests, information on the drug of interest, and specific recommendations for drug or dose changes based on genetic test results. Numerous citations in the literature as well as formal endorsement by the American Society of Health-System Pharmacists (ASHP) are favorable indications of CPIC guideline uptake.[11]

1.1.2 Drug-metabolizing enzymes and transporters

The absorption, distribution, metabolism, and elimination of drugs and other xenobiotics is mediated by drug-metabolizing enzymes and drug transporters. Drug-metabolizing enzymes are responsible for phase I and phase II of drug metabolism. Phase I consists of oxidation, reduction, and hydrolysis reactions, 75% of the reactions catalyzed by cytochrome P450 enzymes. Phase II consists of the conjugation of endogenous hydrophilic moieties by transferase enzymes. Transporters are involved with absorption, distribution, and elimination of drugs and metabolites.

Variability in drug response and toxicity is multifactorial, with sources such as genetics, age, environmental factors, disease states, and drug-drug interactions all playing a role. Of these factors, the genetics of drug-metabolizing enzymes and drug transporters make a large contribution. Studies have found that interindividual variability in function

for drug-metabolizing enzyme and transporter genes is higher than that in genes not related to drug metabolism or transport. In a study that examined protein expression levels in a cohort of 427 liver samples, a 641-fold difference was observed between the highest and lowest expression levels of CYP3A4, a phase I enzyme. For glutathione S transferase A1, a phase II enzyme, a 582-fold difference was observed.[12] Pharmacogenomics offers a means to explain these differences.

1.1.2.1 Phase I Drug Metabolism

Phase I drug metabolism occurs mostly in the liver, with most phase I enzymes located in hepatocyte endoplasmic reticulum membranes. Phase I reactions include oxidation, reduction, and hydrolysis, with oxidation being the most common. These reactions convert lipophilic drugs into more easily excreted hydrophilic compounds. The cytochrome P450 (CYP) family of drug-metabolizing enzymes is responsible for most phase I metabolism.[13]

There are 57 functional human CYP enzymes, grouped into 18 families and 44 subfamilies. Of these, about a dozen enzymes from families 1, 2, and 3 are responsible for the metabolism of 70-80% of all clinically used drugs. There is a wide range of variability in drug-metabolizing enzyme activity between individuals. This variability is often multi-factorial and dependent on factors such as age, sex, and disease state. For some CYP genes, genetic polymorphisms are a major determinant of variability. In the context of drug metabolism, loss-of-function variants can result in decreased drug clearance of the parent compound while gain-of-function variants can result in increased drug clearance.[14]

1.1.2.2 Phase II Drug Metabolism

Phase II metabolism also occurs primarily in the liver. Some phase II enzymes are cytosolic while others are located in hepatocyte endoplasmic reticulum membranes. Phase II reactions are conjugation reactions that increase the hydrophilicity of the parent compound or primary metabolite. Enzymes include UDP-glucuronosyltransferases (UGTs), sulfotransferases (SULTs), N-acetyltransferases (NATs), glutathione S-transferases (GSTs), and methyltransferases, primarily thiopurine S-methyl transferase (TPMT) and catechol O-methyl transferase (COMT).[15] The most prominent phase II enzymes are the UGTs, which glucuronidate 40-70% of all clinically used drugs. Of the four UGT families identified in humans, UGT1 and UGT2 are involved in the glucuronidation of xenobiotics.[15]

1.1.2.3 Drug Transport

Compared to drug-metabolizing enzymes, there are fewer clear pharmacogenomic associations between genetic variation in drug transporter genes and drug efficacy and/or toxicity.[16] Major families of drug transporters include ATP binding cassette (ABC) and solute carrier (SLC and SLCO) transporters.

Key transporters in the ABC family are P-glycoprotein (*MDR1/ABCB1*), breast cancer resistance protein (BCRP/*ABCG2*), multidrug resistance proteins (MRPs/*ABCC*s). Of these transporters, clinical evidence is the best for BCRP pharmacogenetics. BCRP, encoded by *ABCG2*, is an efflux transporter expressed in the liver, intestine, the kidney, and in blood-tissue barriers. Over 180 variants have been identified in the *ABCG2* gene but only a handful have been associated with variation in function, expression, mRNA stability, or clinical outcomes. The most studied variant, 421C>A affects transport rates of statins and possibly tyrosine kinase inhibitors. [17,18] *ABCB1,* the gene that encodes the

efflux transporter P-glycoprotein (P-gp), is highly polymorphic but there is conflicting evidence on whether *ABCB1* genotype has any influence on drug plasma levels or clinical outcomes. [16,18]

SLC drug transporters include but are not limited to organic anion transporters (OATs), organic cation transporters (OCTs), peptide transporters (PEPTs) and organic anion transporting polypeptides (OATPs).[17] A notable pharmacogenetic association exists for hepatic uptake transporter OATP1B1, encoded by gene *SLCO1B1*, and the lipid-lowering agent simvastatin. Simvastatin is a major substrate of OATP1B1, and patients carrying an allele containing the 521T>C variant (*5, *15, or *17) are at higher risk of statin-related myopathy due to increased plasma and peripheral drug levels. A CPIC guideline has been published for *SLCO1B1* with the recommendation to prescribe lower starting doses or alternative statins for individuals heterozygous or homozygous for decreased-function alleles. [19,20]

1.1.3 Applications

Pharmacogenetics was initially implemented in the clinic via single-gene tests ordered if a physician was considering prescribing a drug with an actionable pharmacogenetic association or to guide therapy changes after treatment failure or unacceptable toxicity. This remains the most widespread method of pharmacogenetic testing.[21] The FDA requires pharmacogenetic testing for certain drug/gene pairs and groups such as CPIC provide guidance documents on additional drugs or genes with pharmacogenetic indications.

Some large medical institutions have implemented preemptive pharmacogenomics programs where genetic testing is done on multiple pharmacogenes

when a patient enters into a healthcare system. The data is placed in their electronic medical record and is available if they are ever prescribed a drug with an actionable pharmacogenetic association.[22]

1.1.3.1 FDA-required pharmacogenetic tests

The FDA currently requires genetic testing for 52 drugs and recommends it for 4 more. A test is classified as "required" if the label states that a test "should be performed." In some cases, an FDA-approved companion diagnostic test must be performed before a patient is initiated on a drug. In other cases, the requirement is up to the discretion of the provider. A test is classified as "recommended" if the label states that it "should be considered." [23,24]

The majority of drugs with required genetic testing are kinase inhibitors and monoclonal antibodies used in oncology. Generally, the genetic tests are for determining whether tumor cells contain specific genetic mutations that make them sensitive or resistant to the drug. For example, erlotinib is kinase inhibitor used for treatment of patients with metastatic non-small cell lung cancer whose tumors have specific mutations in the epidermal growth factor receptor (*EGFR*) gene. Only patients who are positive for the mutations, as determined by an FDA-approved companion diagnostic test, may receive erlotinib.[25] An example of a monoclonal antibody requiring genetic testing is trastuzumab. Trastuzumab is used in the treatment of HER2-overexpressing breast cancer and gastric cancer. Prior to initiating therapy, HER2 overexpression must be confirmed by an FDA-approved fluorescence in situ hybridization (FISH) assay to test gene amplification.[26]

Genetic tests may be required when certain genetic polymorphisms are associated with higher risks of adverse reactions. Two of these pharmacogenetic associations are related to mutations in the human leukocyte antigen B (HLA-B) gene. Abacavir, an antiretroviral drug used to treat HIV/AIDS, is associated with severe hypersensitivity reactions in patients with the *HLA-B*5701* allele and testing for the allele is required prior to initiation.[27] Carbamazepine, an anticonvulsant, has been associated with severe dermatologic reactions in patients with the *HLA-B*1502* allele. This allele is most common in patients of Asian descent but mostly absent in other populations. Thus, the FDA states that patients of Asian descent should be screened for *HLA-B*1502* prior to initiating treatment with carbamazepine.[28]

CYP2D6 is the only CYP enzyme with FDA-recommended and –required testing. The FDA-required pharmacogenetic tests are used to guide dose selection for CYP2D6 substrates tetrabenazine, eliglustat, and pimozide. The FDA recommends testing to determine the risk of adverse effects when treating with dextromethorphan/quinidine. For other CYP enzymes, the FDA provides information about pharmacogenetics but does not specifically recommend testing. For example, the clopidogrel drug label explains that efficacy is dependent on conversion to an active metabolite by CYP2C19 and that an alternative platelet inhibitor should be considered in those identified by genotyping as having low CYP2C19 activity.[29]

1.1.3.2 Clinical Pharmacogenomics Implementation Consortium guidelines

To date, the CPIC has published guidelines on 36 drug/gene combinations. Each guideline includes background information on the gene(s) and drug of interest, available genetic test options and how to interpret them, data on how genetic variability is linked to

drug-related phenotypes, and dosage recommendations.[11] The CPIC has also begun providing resources for integration of pharmacogenomic test results into electronic health records.[30]

The guidelines are primarily for drug-metabolizing enzymes, with the majority focused on CYPs. For *CYP2C19*, guidelines are available for SSRIs citalopram, escitalopram, and sertraline, the anti-platelet agent clopidogrel, and the anti-fungal voriconazole. For *CYP2D6,* guidelines are available for SSRIs paroxetine and fluvoxamine, tricyclic antidepressants nortriptyline and desipramine, opioid analgesic codeine, and anti-emetics odansetron and tropisetron. Some guidelines include recommendations based on the diplotypes for multiple genes. Guidelines that include both *CYP2C19* and *CYP2D6* are for tricyclic antidepressants amitriptyline, clomipramine, imipramine, trimipramine, and doxepin. Additional guidelines that include CYPs are published for tacrolimus and *CYP3A5*, phenytoin and *CYP2C9* plus *HLA-B*, and warfarin and *CYP2C9*, *CYP4F2*, and *VKORC1*.[31]

1.1.3.3 Preemptive pharmacogenomics programs

Some healthcare systems have developed preemptive pharmacogenomic testing programs focused on drug/gene pairs with the strongest evidence. Preemptive pharmacogenomics involves genotyping a panel of genes related to drug metabolism and incorporating the data into patients' electronic medical records. If an actionable drug is prescribed, the clinician will receive guidance on how to adjust the dose or choose another drug based on the patient's diplotype. [21,32]

The Pharmacogenomics Research Network (PGRN) Translational Pharmacogenomics Program (TPP) represents a major effort to implement preemptive

pharmacogenomics programs in a variety of clinical sites. The sites that are part of TPP include Mayo Clinic, The Ohio State University, St. Jude Children's Research Hospital, University of Florida, University of Maryland, Vanderbilt University Medical Center, University of Chicago, and Brigham and Women's Hospital. Each site has developed programs relevant to the patient population it serves.[32]

St. Jude Children's Research Hospital has one of the largest preemptive pharmacogenetics programs. In 2011, an array-based genotyping test that interrogates 1,936 variants in 230 genes was implemented as part of routine patient care. As new literature is published, a pharmacogenetics committee decides on which drug/gene pairs to include in the electronic health record (EHR) system. As of January 2014, genetic data was included in the EHR for a total of 1,106 patients. Of these patients, 792 (78%) have at least one actionable drug-metabolizing enzyme diplotype, meaning that a drug or dose change would be necessary. The drug/gene pairs that are included in the EHR represent some of the biggest "success stories" in pharmacogenetics: *TMPT* and thiopurines, *CYP2D6* and codeine, *CYP2C9+VKORC1* and warfarin, *CYP2C19* and clopidogrel, and *UGT1A1* and irinotecan.[33]

Vanderbilt University Medical Center's program, Pharmacogenomic Resource for Enhanced Decisions in Care and Treatment (PREDICT), was initially focused on antiplatelet medications in patients who were either anticipating the need for coronary artery stenting or who had cardiovascular risk scores that suggested they may need antiplatelets in the future. Since coronary artery stenting is often performed as an emergency procedure, having pharmacogenetic data available prior to the event expedites the process of choosing the most efficacious antiplatelet for each specific patient.

Similarly to the methods used at St Jude, genotyping data is stored outside of the EHR until a genetic result is deemed actionable by committee. The first diplotype included in the EHR was *CYP2C19*2/*2*, which is a poor metabolizer diplotype. Clopidogrel is a prodrug that must be metabolized by CYP2C19 to an active metabolite. Poor metabolizers do not produce enough of this active metabolite, resulting in significantly reduced platelet inhibition and an increased risk for adverse cardiovascular events. Patients who are CYP2C19 poor metabolizers should receive an alternative antiplatelet, such as prasugrel or ticagrelor, which is not a CYP2C19-dependent prodrug. [34]As of 2013, the PREDICT program has genotyped 10,000 patients and implemented the following drug/gene pairs in addition to *CYP2C19* and clopidogrel:, *SLCO1B1* and simvastatin, *CYP2C9+VKORC1* and warfarin, *CYP3A5* and tacrolimus, and *TPMT* and thiopurines.[32]

1.2 Cytochromes P450

1.2.1 CYPs responsible for drug metabolism

Cytochrome P450 enzymes are a superfamily of mono-oxygenases responsible for the metabolism of 70 – 80% of all clinically used drugs. They are the primary contributor to phase I drug metabolism via the catalysis of oxidative biotransformation reactions.[14]

CYPs are endoplasmic reticulum-associated membrane-bound enzymes located hepatically, and to a lesser extent, extrahepatically. There are 57 functional P450 genes grouped into 18 families and 44 subfamilies. Of these, families 1, 2, and 3 are responsible for the metabolism of drugs and other xenobiotics. The remaining families have a variety of functions including, but not limited to, the synthesis of steroid hormones, prostaglandins, and bile acids. Substrates include sterols, fatty acids, eicosanoids, and

vitamins. The predominant xenobiotic-metabolizing CYPs in the liver are CYPs 1A2, 2C9, 2C8, 2E1, and 3A4 with 2A6, 2D6, 2B6, 2C19, and 3A5 found less abundantly. [14,35]

The CYP1 family contains CYP1A1, CYP1A2, and CYP1A3. Of the three, CYP1A2 is the most clinically relevant and is responsible for the metabolism for 8.9% of clinically used drugs. It represents ~4-16% of the total hepatic P450 pool while CYP1A1 and CYP1B1 are expressed extrahepatically. Substrates of CYP1A2 include analgesics and antipyretics, antipsychotics, antidepressants, anti-inflammatories, and some cardiovascular drugs. Variation in CYP1A2 activity is primarily due to non-genetic factors such as enzyme induction via tobacco smoke, charred meat, or cruciferous vegetables.[14]

The CYP2 family contains 16 genes physically organized in gene clusters. Genes relevant to drug metabolism are in families CYP2A-E and CYP2J. The CYP2A family, consisting of CYP2A6, CYP2A7, and CYP2A13, makes up ~4% of the hepatic P450 pool. Of the three isoforms, CYP2A6 is the most relevant to drug metabolism and has clinically relevant pharmacogenetic implications. It is the primary enzyme responsible for the metabolism of nicotine into the inactive metabolite cotinine and individuals with low-activity alleles have been shown to have higher plasma nicotine levels. The sole member of the CYP2B family is CYP2B6. It contributes an average of ~2-5% to the total hepatic P450 pool. The gene is highly polymorphic and *CYP2B6* pharmacogenetics are relevant in the context of HIV treatment. CYP2B6 is the primary enzyme responsible for the metabolism of efavirenz. Higher plasma levels of efavirenz have been observed in patients with low-activity alleles, with higher drug exposures associated with neurotoxicity necessitating treatment discontinuation. The only gene in the CYP2E

family is *CYP2E1*. It is expressed at relatively high levels in adult livers and is responsible for the metabolism of ~3% of clinically used drugs. Genetic variation does not make a major contribution to CYP2E1 activity and is instead caused by enzyme induction, inflammation, and certain disease states such as diabetes and liver disease.[14]

The CYP2C family contains CYP2C8, CYP2C9, CYP2C18, and CYP2C19. CYP2C9 is highly expressed in the liver and contributes ~20% of the total hepatic P450 pool. CYP2C8 and CYP2C19 are expressed at 2- and 10-fold lower levels, respectively. Though it is highly expressed, CYP2C18 does not make large contribution to drug metabolism. While genetic variation in *CYP2C9* and *CYP2C19* is clinically relevant, the clinical impact of genetic variation in *CYP2C8* is not clear. CYP2C9 is responsible for the metabolism of ~13% of clinically used drugs including substrates with narrow therapeutic indices, such as warfarin. Genetic variation in *CYP2C9* and *VKORC1* (the pharmacodynamic target of warfarin) can explain 50-60% of the variation in warfarin maintenance dose in Caucasian patients. CYP2C19 is responsible for the metabolism of 6.8% of clinically used drugs and is particularly important in the metabolism of clopidogrel, an anti-platelet agent.[14]

The CYP2D family consists of CYP2D6, which will be discussed separately in section 1.3.

The CYP3 family consists of CYP3A4, CYP3A5, and CYP3A7. The most abundant isoform is CYP3A4. It contributes ~14-24% of the total hepatic P450 pool and is also the major P450 enzyme in intestinal enterocytes. CYP3A4 metabolizes a wide variety of substrates and is responsible for the metabolism of ~30% of clinically used drugs. For these reasons, CYP3A4 is a major contributor to the first-pass metabolism of

many orally administered drugs. Variability in CYP3A4 activity is influenced by multiple factors, such as transcriptional induction by xenobiotics, cytokine-mediated down-regulation secondary to inflammatory responses, and genetic polymorphisms. Clinically, genetic variation in *CYP3A4* has been demonstrated to be relevant to statin response and time to reach appropriate steady-state concentrations of tacrolimus and cyclosporine in transplant recipients.[14]

CYP3A5 shares high sequence similarity with *CYP3A4* and thus similar substrate selectivity. *CYP3A5* is unlike other P450 isoforms in the sense that a non-functional allele is more common in some populations than the *CYP3A5*1* allele. The most common loss of function allele, *CYP3A5*3*, is the found at a frequency of 0.88 – 097 in Caucasian populations and 0.66 – 0.75 in Asians and Hispanics. In people of African descent, the *CYP3A5*3* allele is not as predominant and is found at a frequency of 0.12 – 0.35 in Africans and 0.37 in African Americans.[14]

CYP3A7 is the fetal form of CYP3A and accounts for 50% of total P450 protein content in fetal livers. It has similar substrate specificity as CYP3A4 and CYP3A5. After birth, CYP3A expression shifts from CYP3A7 to CYP3A4, but some individuals express CYP3A7 into adulthood due to a mutation in the promoter region that increases transcription factor binding. The clinical significance of this polymorphism is not well known.[14]

1.2.2 Biochemistry

All cytochrome P450 enzymes are hemoproteins. The presence of iron heme allows the enzyme to utilize oxygen and electrons to ultimately release a molecule of water and transfer an oxygen atom to a substrate.[36]

The overall CYP reaction can be represented as follows:

$$S + O_2 + 2_{e-} + 2H^+ \rightarrow H_2O + SO$$

Cytochrome P450 oxidoreductase (POR) and nicotinamide adenine dinucleotide phosphate (NADPH) are cofactors involved in transfer of electrons to CYPs. POR is a membrane-bound protein that is in the diflavin oxidoreductase enzyme family. In a single polypeptide, it contains one molecule of flavin adenine dinucleotide (FAD) and one molecule of flavin mononucleotide (FMN). NADPH transfers two electrons in the form of a hydride ion to FAD, which then transfers them one by one to FMN which in turn transfers electrons one by one to the heme in the P450 enzyme.[37]

The P450 catalytic cycle begins when a substrate enters the active site and displaces the water molecule bound to the heme. Electron transfer from POR via NADPH causes the reduction of iron(III) to iron(II). Next, molecular oxygen binds to the iron to form a dioxygen adduct. One electron from the iron(II) center and one electron from the oxygen pair form the iron(III)-oxygen bond. The next reduction step, transfer of a second electron from POR via NADPH, is the rate-limiting step for many P450 enzymes. In this step, a negatively charged iron(III)-peroxo complex is formed as an intermediate. The peroxo group is then rapidly protonated to produce a nucleophilic iron(III)-hydroperoxo complex. A second protonation generates a molecule of water and an electrophilic iron(IV)-oxo species that then oxidizes the substrate. The substrate dissociates and the cycle can start over again.[38]

Differences in active site size and conformation confer substrate specificity. P450s in the CYP1A and CYP1B families have narrow, rigid active sites that accommodate planar molecules such as polynuclear aromatic hydrocarbons. CYP3A4 metabolizes a much wider range of substrates and this is reflected in the differences in its active site. CYP3A4 has a large and open active site cavity that allows for more exposure to the heme surface. Also, its active site is not rigid and can expand and contract to accommodate substrates of different size. Of the enzymes in the CYP2 family, CYPs 2E1, 2A6, and 2B6 have small active site cavities and CYPs 2C8, 2C9, 2C19, and 2D6 have large active site cavities. The enzymes with larger active site cavities metabolize a wider range of substrates. Differences within enzyme families confer further substrate specificity. CYP2C19, for example, has two internal cavities. The first is positioned above the surface of the heme where reactive intermediates are formed and the second is an adjacent antechamber that may represent a substrate access channel. CYP2C8, on the other hand, contains only one large cavity. Most P450s have a single substrate-binding site but some can bind more than one substrate molecule, either in the active site or in an allosteric site. [39,40]

1.2.3 Principles of *in vitro* P450 reaction phenotyping

The enzyme kinetic parameters V_{max}, K_m, and CL_{int} can be determined using liver-based *in vitro* P450 reaction assays. The data obtained from these assays can then be used to assess topics such as the clearance of investigational drugs, the effects of drug-drug interactions, and the effects of genetic polymorphisms on drug clearance.[41]

V_{max} refers to the theoretical maximum rate of an enzymatic reaction and occurs when the enzyme is saturated with substrate. The Michaelis constant, K_m, represents the

substrate concentration that yields a reaction velocity that is one half of V_{max}. An enzyme with high affinity for a substrate will have a low K_m value, meaning that 50% of the enzyme molecules will be bound to substrate at a relatively low concentration. Vice versa, an enzyme with a low affinity for a substrate will have a high K_m value. [41,42]

Values for K_m and V_{max} are determined experimentally by incubating enzyme with substrate at a range of concentrations. When the rate of the reaction (v) is plotted against the concentration of substrate ([S]), the result is a hyperbolic curve that can be described using the Michaelis-Menten equation.[43]

$$v = \frac{V_{max} \times [S]}{K_m + [S]}$$

1.2.3.1 In vitro models of drug metabolism

Numerous liver-based methods are used for *in vitro* investigation of hepatic drug metabolism. *In vitro* models include isolated perfused livers, liver slices, isolated or cultured hepatocytes, recombinant enzymes, and human liver microsomes (HLMs). Each model has its own particular advantages and disadvantages.[44] HLMs are the most popular model for studying drug metabolism and thus will be discussed more in-depth in this section.

The use of isolated perfused livers offer the best representation of *in vivo* conditions but use is limited to animal models due to the scarcity of whole human livers. Isolated perfused livers are not generally used to assay CYP-mediated drug metabolism. Due to the presence of functional bile canaliculi, they are most useful for transporter assays when the biliary excretion of a drug is being investigated. Liver slices offer similar

advantages to isolated perfused livers: the presence of drug-metabolizing enzymes, transporters, and functional bile canaliculi. However, it is more difficult to collect bile from liver slices and the presence of necrotic cells may limit drug transporter assays. Furthermore, liver slices are not always practical since CYP expression decreases >50% within 24 hours of preparation and tissue perfusion by the substrate is limited.[44]

Isolated primary hepatocytes are easier to procure and use than whole livers or liver slices and offer a good representation of *in vivo* activity since they reflect natural heterogeneity in CYP expression. A disadvantage is that isolated primary hepatocytes are only viable for 2-4 hours. Cryopreserved human hepatocytes are available commercially and may be cultured for up to four weeks. However, CYP expression decreases with time and can introduce unwanted variability.[44]

Recombinant systems use transgenic human cell lines, bacteria, or yeast to express recombinant CYP enzymes. These systems are advantageous over hepatocytes since the cells are easier to culture and can be designed to express a single CYP, which is useful if a model focused on a single enzyme is desired. A disadvantage is that data obtained from studies using recombinant systems may be difficult to extrapolate to *in vivo* conditions.[44]

HLMs are vesicles of hepatocyte endoplasmic reticulum obtained by differential centrifugation of liver tissue or primary hepatocytes. HLMs contain CYP enzymes as well as other endoplasmic reticulum-associated drug-metabolizing enzymes such UGTs, flavin monooxygenases, carboxyl esterases, and epoxide hydrolase. Advantages to using HLMs for studying CYPs *in vitro* is that they are cheap, are relatively easy to use, can be frozen and stored for years, and allow study of inter-individual variability. A

disadvantage to using HLMs is that *in vitro-in vivo* extrapolation is limited by the fact that contributions from drug-metabolizing enzymes not present in the endoplasmic reticulum fraction are not accounted for. Furthermore, cytosolic cofactors involved in CYP reactions must be added to the system exogenously.[44,45]

HLMs are available commercially as pooled preparations or through liver banks as single-donor preparations and the choice to use one or the other depends on the goals of a particular study. Pooled products are useful when the goal of a study is to reduce interindividual variability and assess how a population as a whole metabolizes a xenobiotic. HLM pools can be designed to be representative of a broad population, such as Americans as a whole, or subpopulations, such as African Americans. They can also be designed to include HLMs from, for example, only donors genotyped to be CYP2D6 poor metabolizers. In the latter case, a more specific population is being studied, but pooling several donors theoretically results in a HLM product that is representative of the average. Single-donor HLMs are used when the goal of a study is to investigate interindividual variability in drug-metabolizing enzymes. In these types of studies, deviations in enzyme activity from average population levels are of particular interest so pooling several samples would be counterproductive. [45,46]

1.2.3.2 Human liver microsome incubations

In vitro assays with HLMs involve incubating the HLMs with a probe substrate, NADPH, and a phosphate buffer at 37 °C for a specified amount of time in a shaking water bath. After stopping the reaction by adding a protein precipitating reagent such as ice cold acetonitrile, either metabolite formation or substrate depletion is quantified using methods such as LC-MS/MS.[45]

A number of methodological considerations must be made when carrying out HLM incubations. NADPH should not be a rate-limiting reagent and should be used at saturating concentrations of 0.5 – 1 mM. The incubation should be carried out in physiologically relevant conditions through the use of phosphate buffers at a concentration of 50 – 100 mM, a pH of 7.4, and a temperature of 37 °C.[43]

The ability to calculate intrinsic clearance requires the reaction to be run under linear conditions. Methodological factors that determine linearity include enzyme concentration, substrate concentration, and incubation time. Enzyme concentrations must be much less than the substrate concentration to ensure that the reaction rate does not slow due to reduced substrate availability or through product inhibition. A generally accepted number is a 5% reduction in substrate during the course of the assay. At the same time, substrate concentration must be low enough to ensure that $[S] << K_m$. Incubation time should be selected in a way that results in negligible product formation relative to substrate concentrations. Incubation time is of further concern since short incubations can be impractical and long incubation times can cause enzyme and substrate degradation.[43]

When a K_m value for a specific substrate can be obtained from the literature, optimization of incubation conditions is relatively simple. Ideally, 8-10 different substrate concentrations that span the range of one tenth to ten-fold of the K_m value is ideal, though one fifth to five-fold is suitable if there are limitations related to solubility or analytical methods. While keeping the substrate concentrations constant, enzyme concentration and incubation time can be varied to fit the considerations discussed above. From these data points, the kinetic data can be plotted and the most appropriate substrate concentration,

enzyme concentration, and incubation time can be selected for further assays.[43]

Two distinct experimental approaches, metabolite formation and substrate depletion, are used for kinetic experiments. Metabolite formation is the traditional approach used to determine CL_{int}. This approach uses short incubation times and low enzyme concentrations to fulfill the assumption of minimal substrate consumption. The disadvantage to the metabolite formation approach is that the metabolic pathway must be known. When the metabolic pathway of a drug is not known, the substrate depletion approach may be useful. However, with this approach, at least 20% of the substrate must be metabolized, requiring longer incubation times and higher enzyme concentrations. This increases the likelihood of product inhibition and degradation of the substrate and/or enzyme. Furthermore, studies to determine linearity are not often performed so CL_{int} values obtained using a substrate depletion approach may have poor *in vivo* predictive utility.[47]

1.2.4 Allele nomenclature

An allele is a variant form of a gene. Humans are diploid organisms, meaning they have two alleles at each genetic locus.[48] Some alleles consist of a single nucleotide variant (SNV) occurring in either a coding (exonic) or non-coding (intronic) region of the gene. SNVs can also occur in the promoter region or in upstream or downstream untranslated regions flanking the gene. Alleles can be defined by the presence of insertions/deletions (indels), splicing variations, or structural variation like gene deletions, duplications, and recombinations. Alleles can also represent haplotypes, or a cluster of SNVs that are inherited together.[49]

Each P450 gene has its own allele nomenclature that is based off of historical naming conventions. The Pharmacogene Variation Consortium (PharmVar) has sought to standardize P450 naming with a star allele nomenclature system. In this system, a common allele conferring normal activity is defined as the reference allele for a particular CYP gene and is designated as the **1* allele (e.g. *CYP2D6*1).* Variant alleles are compared to the **1* allele and given their own star name (e.g. **2, *3,* etc.) upon approval of the Human Cytochrome P450 (CYP) Allele Nomenclature Committee. Some alleles have one or more suballeles (e.g. **2A, *2B, *2C)* if they share a common defining polymorphism, such as a SNV resulting in a loss of function, but vary in any additional polymorphisms that do not appreciably change protein function.[50]

An individual's diplotype represents both alleles at the gene locus of interest. An example of a diplotype would be **1/*2*, which indicates that the individual has one copy of the **1* allele and one copy of the **2* allele.

1.2.5 Translating CYP diplotype to phenotype

A barrier to the wide-scale implementation of pharmacogenomics has been the ability to translate of genetic data into information that is clinically useful. Translational research is focused on the "bench-to-bedside" process of moving knowledge gained through basic science research into medical practice.[51] With pharmacogenomics, this involves using genotyping data to make predictions about the way an individual will respond to a drug and selecting the drug and dose that is most appropriate for them.

In the context of pharmacogenomics, "phenotype" generally refers to efficacy or toxicity and can be observed in multiple ways and defined qualitatively and quantitatively. For example, efficacy may be defined more qualitatively as clinical

response or quantitatively as a particular plasma drug level. Likewise, toxicity can be defined in more qualitative clinical terms or it may be defined using a quantitative measure such as platelet count.

The link between genotype and phenotype is often most obvious with severe adverse drug reactions. One possible cause of an adverse drug reaction is genetic variation in drug-metabolizing enzymes or transporters. For example, statin-induced myopathy is 17x more likely to occur in individuals who are homozygous for a loss-of-function variant in the *SLCO1B1* drug transporter gene. Another cause is genetic variation in genes involved with immune response, such as the increased risk of severe dermatologic reactions to abacavir in those with the *HLA-B*5701* allele. A third cause is variation in genes involved with the target or mechanism of a drug, such as increased risk of warfarin-induced skin necrosis in individuals with genetic protein C deficiency.[2]

Efficacy or CYP enzyme activity are difficult phenotype to link directly to genotype because they are also affected by environmental factors, demographic factors such as age or sex, and interindividual variability in additional genes related to the drug's metabolism or mechanism of action. Because phenotypes such as efficacy or activity are complex traits, the contribution of genetics is defined by the percentage of variation in phenotype predicted. Genetic variants that explain 20 – 40% of phenotypic variance are considered to have good discriminative power.[2]

A classification system is used in which predicted CYP enzyme activity phenotype is described using the terms normal metabolizer, intermediate metabolizer, poor metabolizer, and ultrarapid metabolizer.[52]

Normal metabolizers (NMs) generally have a "normal" amount of enzyme activity due to two copies of a fully functional allele. Intermediate metabolizers (IMs) have reduced enzyme activity relative to normal metabolizers. This may be due to two copies of a partially defective allele, a combination of a normal allele and a partially defective or null allele, or a combination of a partially defective allele and null allele. Individuals who are poor metabolizers (PMs) have no or very little detectable enzyme activity due two copies of a non-functional allele. These non-functional alleles may be due to gene deletions, insertions/deletions, splicing defects, or other polymorphisms that inhibit gene translation or cause a loss of protein function. Ultrarapid metabolizers (UMs) have one or more duplications of functional genes resulting in higher enzyme activity levels than what is typically seen in NMs.

While this system has provided some consistency in prescribing a phenotype to an individual from genetic data, there can still be large interindividual variability within each phenotype class. Furthermore, there may be overlap between phenotype classes. For example, an individual who is predicted to be a normal metabolizer may have lower activity of the CYP of interest than someone who is predicted to be an intermediate metabolizer. This may be due to confounding factors such as concomitant medication use or age- or sex-related differences in enzyme expression.[14] Another possibility is that the NM individual may have an additional uncharacterized genetic polymorphism that results in a loss or gain of enzyme activity.

1.3 CYP2D6

1.3.1 Clinical relevance

CYP2D6 is responsible for the metabolism of 25% of all clinically used drugs from a wide range of drug classes including antiarrhythmics, tricyclic antidepressants, selective serotonin reuptake inhibitors, antipsychotics, beta-blockers, selective estrogen receptor modulators, and opioid analgesics.[14]

To date, the CPIC has published guidelines that involve 12 different CYP2D6 substrates. Guidelines that focus solely on CYP2D6 are for codeine, anti-emetics tropisetron and ondansetron, tricyclic antidepressants (TCAs) desipramine and nortriptyline, and selective serotonin reuptake inhibitors (SSRIs) fluvoxamine and paroxetine. The guidelines for the TCAs amitriptyline, clomipramine, doxepin, imipramine, and trimipramine also consider the diplotypes of both CYP2D6 and CYP2C19 when making recommendations.

While the CPIC guidelines highlight drug/gene pairs with strong pharmacogenetic associations, it is not always clear-cut. The utility of *CYP2D6* pharmacogenetic testing was questioned when two large clinical studies reported that *CYP2D6* diplotype did not predict clinical benefit in women with breast cancer being treated with tamoxifen. [53,54] This was controversial because previous studies had found an association between *CYP2D6* diplotype and plasma levels of endoxifen, an active metabolite formed by *CYP2D6.* Adding to the controversy were issues with DNA quality and genotyping. A major criticism of these studies was that the tumor sample DNA used for genotyping was inappropriate and that results suggested undetected *CYP2D6* deletion due to loss of heterozygosity. [54,55] Another issue is that the authors only tested a limited number of *CYP2D6* variants and did not assess copy number variation to detect deletions or

duplications. Despite these flaws, the studies have had widespread effects on clinical practice. For example, the National Comprehensive Cancer Network references the studies and states that they do not recommend *CYP2D6* genotyping in patients treated with tamoxifen, due to "limited and conflicting evidence."[56]

1.3.1.1 Codeine

Variation in codeine metabolism represents a classic example of the clinical relevance of *CYP2D6* pharmacogenetics. Codeine is considered a prodrug; while codeine has mild analgesic effects, its active metabolite, morphine has a 200-fold greater affinity for the μ-opioid receptor. The formation of morphine is dependent upon the CYP2D6-mediated O-demethylation of codeine. In poor metabolizers, the percentage of codeine that is metabolized to morphine is lower than what is observed in normal metabolizers. Studies have shown that poor metabolizers experience a decrease in analgesia as well as a decreased incidence of side effects such as constipation. In ultrarapid metabolizers, a higher percentage of codeine is metabolized into morphine than what is observed in normal metabolizers. This represents the potential for more severe adverse effects as increased morphine concentrations can result in respiratory depression or cardiac arrest.

The potential for serious adverse effects was fully realized in 2005 with the death of a 13-day-old infant.[57] His mother had been prescribed codeine for post-partum pain and took it while breastfeeding, consistent with the American Academy of Pediatrics listing codeine as being compatible with breastfeeding. Post-mortem analysis of the infant found a morphine serum concentration of 70 mg/mL, much higher than the typical value for breastfed infants of 0.2 – 2 ng/mL. Breastmilk that the mother had stored from postpartum day 10 had a morphine concentration of 87 ng/mL. Typical values observed

in breastmilk from women taking similar codeine doses range from 1.9 – 20.5 ng/mL. The mother was genotyped for *CYP2D6* and was found to be an ultrarapid metabolizer, explaining the higher than normal concentration of morphine in her breastmilk.[57]

Following this case report, the FDA changed the codeine product label to include a boxed warning cautioning against its use in children and information on the risks of opioid toxicity in infants breastfed by mothers who are CYP2D6 ultrarapid metabolizers. The label mentions that there is a pharmacogenetic test available but that they are not routinely performed.[58] The CPIC has since issued a guideline on codeine and *CYP2D6* for use in the general population, recommending that codeine be avoided in ultrarapid metabolizers and poor metabolizers in favor of alternate analgesics such as morphine or non-opioid analgesics.[59]

1.3.1.2 Tricyclic antidepressants

Tricyclic antidepressants have largely fallen out of favor due to the high incidence of side effects but still have roles in the treatment of refractory depression and pain. Certain tricyclic antidepressants (amitriptyline, clomipramine, doxepin, imipramine, and trimipramine) are demethylated by CYP2C19 to active metabolites. CYP2D6 then hydroxylates them to less active metabolites. Desipramine and nortriptyline are not metabolized by CYP2C19 but are hydroxylated by CYP2D6 to less active metabolites. The CPIC guideline gives recommendations based on both *CYP2C19* and *CYP2D6* diplotype, but only the *CYP2D6* recommendations will be discussed presently.

In CYP2D6 poor metabolizers, the metabolism of tricyclic antidepressants to less active metabolites is reduced relative to normal metabolizers. The risk for adverse effects is increased due to high plasma concentrations of more active metabolites of the drugs.

The CPIC recommendation is to avoid the use of TCAs in poor metabolizers. If a TCA must be used, the recommendation is to use a starting dose that is 50% lower than what is recommended in the drug label and then to use therapeutic drug monitoring to guide dose adjustments. In intermediate metabolizers, metabolism of TCAs is reduced, but not to the extent that is observed in poor metabolizers. The CPIC recommends a 25% reduction in starting dose and therapeutic drug monitoring. For ultrarapid metabolizers, the concern is that TCAs will be metabolized at an increased rate, resulting in a loss of therapeutic efficacy. In this case, the CPIC recommends avoiding TCAs or increasing the starting dose and using therapeutic drug monitoring to guide dose adjustments.[60]

1.3.1.3 Selective serotonin reuptake inhibitors

Selective serotonin reuptake inhibitors have supplanted tricyclic antidepressants as first-line therapy for the treatment of major depressive disorder. Like TCAs, SSRIs are metabolized by CYP2C19 and CYP2D6. The efficacy and safety of SSRIs vary among individuals; up to 50% will fail initial SSRI therapy and 25,000 patients visit emergency departments every year due to SSRI-related adverse events. Some of this variability in response is due to *CYP2C19* and *CYP2D6* pharmacogenetics.

Paroxetine and fluvoxamine are metabolized to inactive metabolites by CYP2D6. Ultrarapid metabolizers are of the greatest concern, with multiple studies demonstrating there is at ultrarapid metabolizers have low plasma paroxetine concentrations, putting them at risk for treatment failure. The CPIC recommendation is that ultrarapid metabolizers be treated with an alternative drug that is not metabolized by CYP2D6. There is less evidence of treatment failure in ultrarapid metabolizers with fluvoxamine, so the CPIC does not make a specific recommendation. Similarly, even though

intermediate metabolizers would be expected to have decreased metabolism and thus increased risk for adverse effects, there is not enough evidence to make recommendations for adjustment to paroxetine or fluvoxamine therapy. Poor metabolizers have greater exposure to paroxetine and fluvoxamine, potentially increasing the risk for adverse effects. Per the FDA, the fluvoxamine drug label includes a precaution about use in CYP2D6 poor metabolizers. Based on data from pharmacokinetic modeling, the CPIC recommends a 50% reduction in starting dose for paroxetine and a 25-50% reduction for fluvoxamine.[61]

1.3.1.4 Emerging areas

The CPIC ranks drug/gene pairs based on the level of supporting evidence in the literature and prioritizes those with the highest level of evidence for new guidelines. Opioid analgesics oxycodone and tramadol and selective estrogen receptor modulator tamoxifen have an evidence rating of "A" and will likely be the next areas of focus for guideline publication. Drugs with an evidence level of B are considered next. These are primarily psychiatric drugs and include antipsychotics risperidone, brexpiprazole, aripiprazole, pimozide, and mirtazapine, and antidepressants sertraline, venlafaxine, and vortioxetine.[62]

The combination of dextromethorphan and quinidine is marketed for the treatment of pseudobulbar affect, a condition characterized by involuntary laughing and/or crying. Dextromethorphan, a CYP2D6 substrate, is responsible for the therapeutic effect and quinidine, a CYP2D6 inhibitor, is used to increase dextromethorphan plasma levels. Quinidine has no effect on dextromethorphan levels in CYP2D6 poor metabolizers so

genotyping may spare patients from the potentially serious adverse effects associated with quinidine, such as thrombocytopenia and *torsades de pointes.*[63]

1.3.2 CYP2D6 gene locus

The *CYP2D6* gene is located in a segment on chromosome 22q13.2 that also contains the pseudogenes *CYP2D7* and *CYP2D8*.[64] *CYP2D6* itself is comprised of approximately 5350 bp. All of the genes contain nine exons. Both pseudogenes have similar nucleotide sequences to *CYP2D6,* with 92-97% homology across introns and exons.[65]

CYP2D8 is downstream from *CYP2D6* and does not encode a functional protein product due to several insertions, deletions, and sequence variations that code for stop codons. No mRNA has been observed in humans. The *CYP2D7* pseudogene is downstream from *CYP2D8.* It contains a nucleotide insertion in the first exon that results frameshift responsible for a premature stop codon and thus a nonfunctional protein.[66] Several intronic and exonic regions of *CYP2D7* are identical to *CYP2D6*, as well as the 3'-ends and downstream repetitive regions. These similarities can result in gene rearrangements that produce deletions, duplications, and hybrids of *CYP2D6* and *CYP2D7*.[64]

1.3.3 CYP2D6 nomenclature

The reference allele used to define the *CYP2D6*1* sequence is GenBank entry AY545216, submitted in 2005 by Gaedigk et al.[67] This sequence supersedes GenBank entry M33388, the first *CYP2D6* sequence published in 1989 by Kimura et al.[68] While AY545216 corrects sequencing errors found in the previous reference sequence, M33388 is still used by convention to define nucleotide positions used in official allele definitions.

Numbering starts with 1 at the M33388 sequence ATG start codon and has been numbering guidelines have updated to reflect corrections to the original sequencing errors.[64]

Currently, there are 113 primary star alleles defined for CYP2D6. Some of these alleles have multiple suballeles. For example, *CYP2D6*4* is defined by the presence of 1846G>A, a SNV that introduces an alternative splice site leading to nonfunctional truncated CYP2D6 protein.[69] There are currently fifteen **4* suballeles defined in the nomenclature: *CYP2D6*4A-*4P* and *CYP2D6*4X2*, which represents a duplication of the *CYP2D6*4* allele. All of these alleles have the 1846G>A SNV in common but vary in what additional SNVs they contain.[70] The additional SNVs do not change the functional consequence; every *CYP2D6*4* suballele results in nonfunctional CYP2D6 protein.

Structural variants are not always obvious by their nomenclature. *CYP2D6*5* represents the deletion of the *CYP2D6* gene. Hybrid alleles, which are combinations of *CYP2D6* and *CYP2D7*, are given star designations, such as **13*. Duplications are described by listing the allele that is duplicated and the number of copies present on one chromosome. The generic nomenclature for a duplicated allele has "*xN"* after a star allele, e.g. **1xN.* If there were 4 copies of the **1* allele present, it would be designated as **1x4*. Tandem alleles occur when two different alleles are found on the same chromosome and are described as allele 1 + allele 2, e.g. **68* + **4.*

1.3.4 Activity score

The activity score system was introduced by Gaedigk et al. in 2008 as a way to simplify and standardize the clinical translation of *CYP2D6* diplotype data.[52] It has been

adopted by the Clinical Pharmacogenomics Implementation Consortium (CPIC) as the preferred method of translating diplotype into phenotype in *CYP2D6* guidelines.[64]

One of the more common ways to translate drug-metabolizing enzyme diplotype data into a phenotype prediction is with system that presents predicted phenotypes as poor metabolizers, intermediate metabolizers, normal metabolizers, or ultrarapid metabolizers. This method is challenging for *CYP2D6* due to the large number of alleles and potential complexity of particular allele combinations.

With the activity score system, a score of 0 is given to nonfunctional alleles, 0.5 to reduced function alleles, or 1 to functional alleles with activity comparable to the *CYP2D6*1* reference allele. Gene duplications are assigned a score that is a multiple of single gene score (e.g. a duplication of the **41* reduced function allele (**41x2*) would have a score of 1 since **41* has an activity score of 0.5). Diplotype activity scores are thus the sum of the individual allele scores that comprise the diplotype. Possible diplotype scores are 0, 0.5, 1, 1.5, 2, and >2.[64] The activity scores are then translated into predicted phenotypes of poor metabolizers, intermediate metabolizers, normal metabolizers, or ultrarapid metabolizers. Activity scores are qualitative and the scale does not imply that an allele with a score of 0.5 has 50% of the activity of an allele with an activity score of 1.

1.3.5 Relevant *CYP2D6* alleles

*CYP2D6*1* is the reference allele used to define a "wildtype" sequence that encodes a CYP2D6 enzyme with normal activity. An allele is designated as *CYP2D6*1* if it does not include any of the other SNVs found in the alleles indexed by the CYP2D6

allele nomenclature database.[70] *CYP2D6*1* is usually the most common allele found in European and African populations.

*CYP2D6*1xN* refers to the duplication of the **1* allele, with N indicating how many copies of *CYP2D6*1* are present on a chromosome. Multiple copies of the **1* allele results in an ultrarapid metabolizer phenotype.[71]

*CYP2D6*2* is defined by the presence of two missense SNVs, 2850C>T and 4180G>C, that cause amino acid changes R296C and S486T. Functionally, *CYP2D6*2* has slightly reduced activity compared to *CYP2D6*1* but the difference is not considered to be significant, resulting in *CYP2D6*2* having an NM designation.[72]

*CYP2D6*3* is defined by a deletion, 2549delA, that causes a frameshift mutation. The frameshift causes premature truncation of the CYP2D6 protein, resulting in a complete loss of function and a PM designation.[73]

*CYP2D6*4* is defined by the presence of the intronic SNV 1846G>A, which causes a disruption of the splice acceptor site in intron 3.[74] *CYP2D6*4* is nonfunctional and designated as a PM.

*CYP2D6*4xN* refers to the duplication of the **4* alelle, with N indicating how many copies of *CYP2D6*4* are present on a chromosome. Since **4* is nonfunctional, multiple copies do not result in increased CYP2D6 activity, so the allele is designated as a PM.[71]

*CYP2D6*4N+*4* consists of the hybrid allele **4N* in tandem with **4N*. Because both **4N* and **4* are nonfunctional, it is a PM allele.[75]

*CYP2D6*5* is the total deletion of the *CYP2D6* gene caused by homologous recombination of the rep 7 and rep 6 regions and looping out of the gene.[76] Since no

functional protein is produced from a deleted gene, *CYP2D6*5* has the designation of PM. The current PGRNSeq platform is not designed to detect the *CYP2D6*5* allele but a novel bioinformatics tool, Stargazer (discussed further in Chapter 2), has been developed to determine *CYP2D6* copy number using PGRNSeq read depth data.

*CYP2D6*6* is defined by a deletion of a single nucleotide in exon 3, 1707delT, that causes a frameshift. The frameshift leads to the generation of a premature stop codon located one codon after the deletion. The truncated CYP2D6 protein is nonfunctional and results in a PM designation.[77]

*CYP2D6*9* is defined by the deletion of three nucleotides, 2615_2617delAAG, resulting in the deletion of a single amino acid, K281del. The *CYP2D6*9* allele confers reduced function and *in vitro* studies suggest that this is due to reduced expression rather than changes in catalytic activity.[78]

*CYP2D6*10* is defined by the SNVs 100C>T, 1661G>C, and 4180G>C and has reduced function. The 100C>T variant causes a deleterious protein change, P24S, that has been associated with reduced CYP2D6 expression and catalytic activity. [79]

*CYP2D6*17* is defined by the presence of four SNVS, 1023C>T, 1661G>C, 2850C>T, and 4180G>C. *CYP2D6*17* is a reduced function allele. This effect appears to be due to substrate-specific changes in K_m. [80]

*CYP2D6*20* contains an insertion, 1973_1974insG, that causes a frameshift, generating a premature stop codon. The truncated CYP2D6 protein is nonfunctional and results in a PM designation.[81]

*CYP2D6*29* is defined by the presence of the SNVs 1659G>A and 3183G>A, which cause protein changes V136I and V338M. The *CYP2D6*29* allele encodes an

enzyme with reduced function, possibly due to changes in K_m that are substrate-specific.[82]

*CYP2D6*33* is defined by the presence of a single SNV, 2483G>T, which causes the protein change A237S. *CYP2D6*33* has normal activity.[70]

*CYP2D6*35* is defined by the presence of SNVs 31G>A, 2850C>T, and 4180G>C. The 31G>A SNV results in the protein change V11M. *CYP2D6*35* confers normal protein activity.[70]

*CYP2D6*36+*10* is a hybrid tandem allele, meaning that two *CYP2D6* alleles, **36* and **10*, exist together on the same chromosome. **36* is nonfunctional but **10* has intermediate activity, so the overall allele is designated as an IM.[71]

*CYP2D6*39* is defined by the presence of 1661G>C and 4180G>C in the absence of 2850C>T. It has normal activity.[70]

*CYP2D6*41* is defined by the presence of intronic SNV 2988G>A, which causes a splicing defect, in combination with 1661G>C, 2850C>T, and 4180G>C. *CYP2D6*41* is a reduced function allele. The hypothesized mechanism is reduced protein expression secondary to the splicing defect.[83]

*CYP2D6*41x2* refers to the duplication of the **4* alelle, with two copies of **41* on the same chromosome. Since *41 has reduced function, two copies result in roughly normal CYP2D6 activity, designating it as an NM.[71]

*CYP2D6*43* is defined by the presence of a single SNV, 77G>A, which results in the protein change R26H. This allele confers normal function.[70]

*CYP2D6*59* is defined by the presence of the SNVs 2291G>A and 2939G>A in combination with 1661G>C, 2850C>T, and 4180G>C. *CYP2D6*59* is a reduced function

allele. Of the SNVs unique to *CYP2D6*59*, 2291G>A is found in intron 4 and it is not clear if it contributes to the intermediate metabolizer phenotype. The second SNV, 2939G>A, causes a synonymous change in exon 6. The mechanism has not been fully elucidated, but it is hypothesized that the 2939G>A SNV may cause splicing defects or that it may cause changes in secondary mRNA structure.[84]

*CYP2D6*68+*4* is a hybrid tandem allele, meaning that two *CYP2D6* alleles, **68* and **4,* exist together on the same chromosome. Both alleles are nonfunctional, so the overall allele is designated as a PM.[71]

1.3.6 Other sources of CYP2D6 variation

1.3.6.1 Ethnicity

CYP2D6 allele frequencies vary across world populations. Europeans have a relatively high frequency of individuals with a poor metabolizer phenotype. This is largely due to the loss of function **4* occurring at a frequency of 18% in people of European descent. It is lowest in East Asians with a frequency of 0.6% Intermediate metabolizer phenotypes are most common in East Asian, African, and Middle Eastern populations. The intermediate function allele, **10*, is most common in East Asians and occurs at a frequency of 45%. The **17* and **29* alleles are most common in African populations and occur at frequencies of 20% and 9%, respectively. Hybrid genes appear to occur more often in Africans, as well. The *41 allele is most common in Middle Eastern populations, occurring at a frequency of 20%. Ultrarapid metabolizer phenotypes are found most often in Oceanian and Middle Eastern populations due to the high frequency of alleles containing gene duplications.[85-87]

1.3.6.2 Age and sex

Data on the effect of age and sex on CYP2D6 activity are conflicting and vary depending on the probe substrate used.[88]

1.3.6.3 Drug-drug interactions

CYP2D6 is subject to inhibition by a variety of drugs, potentially resulting in drug-drug interactions. Strong inhibitors, which increase the AUC of a second CYP2D6-metabolized substrate by ≥5-fold, include bupropion, fluoxetine, paroxetine, quinidine, and terbinafine. Moderate inhibitors, which increase the AUC of a second substrate by ≥2 to <5-fold, include cimetidine, cinacalcet, duloxetine, fluvoxamine, and mirabegron.[89] Paroxetine and cimetidine are mechanism-based inhibitors and the others are reversible.[71]

1.3.6.4 Liver disease

The effects of liver disease on CYP2D6 activity vary with etiology. In general, CYP2D6 is not as sensitive to liver disease as other CYPs such as CYP1A2, 2C19, and 3A4/5. Patients with chronic hepatitis C have been reported to have 2.6x less CYP2D6 activity versus those without hepatitis C. Anti-CYP2D6 antibodies are present in those with autoimmune hepatitis and have been identified in some patients with chronic hepatitis C. In one study, patients who produced these antibodies had CYP2D6 activity levels 80% lower than those who did not produce antibodies.[71]

1.3.6.5 Cytochrome P450 Oxidoreductase

Cytochrome P450 oxidoreductase (POR) transfers electrons from NADPH to CYP enzymes, a rate-limiting step in CYP-mediated drug metabolism. The *POR* gene is polymorphic and certain variants cause a syndrome consisting of skeletal dysplasia, impaired steroidogenesis, and genital ambiguity. Other variants do not cause a disease

phenotype but have been associated with decreased CYP enzyme function. A POR variant may only affect the activity of certain CYP enzymes and the effect may be substrate-specific. Because POR is required for CYP enzyme function, genetic variation in the *POR* gene may explain some of the variability in CYP activity not explained by CYP genetics. A study examining the effects of various recombinant *POR* variants on CYP2D6 activity found that POR with common variant A503V decreased CYP2D6 activity by 40-50% when using dextromethorphan and bufurolol as probes.[90]

1.3.6.6 Aldo-keto reductase 1D1

Aldo-keto reductase 1D1 (AKR1D1) is involved in the bile acid homeostasis and is hypothesized to regulate CYPs through activation of nuclear receptors that transcriptionally regulate CYP expression. Bile acids are ligands of nuclear receptors PXR and CAR, which induce the transcription of *CYP3A4*, *CYP2C9,* and *CYP2C19.* Since *CYP2D6* is generally considered to be non-inducible, this mechanism of CYP regulation should not affect CYP2D6 activity. However, it is possible that there are additional pathways that have not yet been described.[91]

1.3.7 *CYP2D6* genotyping methods

CYP2D6 is technically difficult to sequence due to a large number of variants, copy number variation, the presence of hybrid and tandem alleles, repeat regions, and high sequence homology with the *CYP2D7* pseudogene.[92]

Targeted genotyping is a method used to test for the presence of specific variants using methods such as TaqMan PCR assays. This method can be problematic for *CYP2D6* because rare variants that alter enzyme activity will not be detected, nor will more common variants that were not included on the genotyping panel. When no variants

are detected, allele assignment "defaults" to *CYP2D6*1* and individuals may erroneously be predicted to be normal metabolizers. Also, targeted genotyping will not provide information on copy number variation unless it is specifically assessed using quantitative real-time PCR. Results of this type of copy number variation assay may be ambiguous since it only reports a generic duplication signal. In diplotypes such as *1/*4, either allele may be duplicated with different effects on enzyme function. *CYP2D6*1* is a functional allele, so a duplication would cause increased function; on the other hand, *CYP2D6*4* is non-functional so a duplication would not change CYP2D6 phenotype.[92]

Sequencing allows for the detection of uncommon or rare variants but has its own unique technological challenges. Short-read sequencing allows for read lengths of up to 300 bp and is the most common sequencing method used commercially. Repetitive regions and pseudogenes are issues for short-read sequencing because the reads may not be long enough to distinguish between regions of the gene locus. Long-red sequencing platforms currently have average read lengths of approximately 10 kb, offering a large improvement over short-read sequencing.[92]

Targeted exome sequencing of selected pharmacogenes, such as with the PGRNSeq platform, offers an approach that can be conducted for less time and money than whole genome or exome sequencing. By focusing on selected genes, targets can be sequenced with deep coverage at relatively high throughput.[93] Furthermore, read depth data can be used to assess *CYP2D6* copy number and for the presence complicated structures such as hybrid genes.

1.4 Specific aims

Aim 1: Assign *CYP2D6* haplotypes to human liver microsome samples from a panel of livers and determine expected CYP2D6 activity scores. Use sequencing data provided by the PGRNSeq platform to assign common *CYP2D6* haplotypes and identify any novel or rare coding variants.

Aim 2: Perform an *in vitro* probe study to determine CYP2D6 activity in human liver microsome samples. Identify SNVs that may explain phenotypic outliers. Use CYP2D6 probe drugs dextromethorphan and metoprolol in human liver microsome (HLM) incubations to estimate CYP2D6 activity.

Aim 3: Incorporate Stargazer data on structural and copy number variation to evaluate the change in the percent variation in CYP2D6 activity predicted by activity score. Reassign diplotypes using Stargazer data and determine if correlation of *CYP2D6* activity with activity score is improved.

2 A multifactorial cytochrome P450 2D6 genotype-phenotype prediction model to improve precision of clinical pharmacogenomic tests

2.1 Introduction

Cytochrome P450 2D6 (CYP2D6) contributes to the metabolism of up to 25% of medications from widely-used classes such as beta-blockers, antiarrhythmics, selective serotonin reuptake inhibitors, tricyclic antidepressants, antipsychotics, and opioid analgesics.[94] A high degree of interindividual variability in CYP2D6 activity has been observed and much of this can be attributed to variation in the highly polymorphic *CYP2D6* gene.[95]

The combination of a wide range of clinically-relevant substrates and a well-studied link between genetics and activity has made *CYP2D6* a logical candidate for clinical pharmacogenetic testing. Currently, information on *CYP2D6* pharmacogenetics is included in the labeling for 52 FDA-approved drugs.[96] To provide guidance on implementation of pharmacogenomics testing in the clinic, groups such as the Clinical Pharmacogenetics Implementation Consortium (CPIC) have published guidelines for 12 gene-drug combinations from 4 therapeutic classes that are metabolized by CYP2D6.[97]

The *CYP2D6* gene is notoriously difficult to interrogate due to challenges imposed by the complexity of the *CYP2D6* gene locus. To date, over 100 *CYP2D6* alleles have been defined, many of which are comprised of multiple variants. Techniques such as targeted genotyping are of limited utility for *CYP2D6* because logistics limit the number of many possible variants to test. Such strategies may miss clinically relevant variants that were present in an individual but not included in the platform. Because of

the way many *CYP2D6* haplotypes build on a **1* or **2*-like "background," an individual may have a deleterious variant, but if it is not tested for, their diplotype will default to **1* or **2,* predicting normal CYP2D6 activity when it is actually reduced or absent. [64,92] This is particularly problematic in diverse populations, where data is lacked about relevant functional variation; if functional variants are missed in these populations, a default **1* or **2* variant would incorrectly assign individual phenotypes. [85,92]

In addition to sequence variation, structural variation is an important consideration in assigning *CYP2D6* diplotypes. The *CYP2D6* gene locus has a high degree of structural variation that results in not only copy number variation (CNV) like gene deletions and duplications, but also gene rearrangements between *CYP2D6* and the adjacent pseudogene *CYP2D7* with high homology to *CYP2D6*, which lead to *CYP2D6-CYP2D7* hybrids. Structural variation is not uncommon: the *CYP2D6*5* allele, a deletion of the *CYP2D6* gene, occurs at a frequency of 2 – 6% worldwide, while duplications that cause increased function, such as *CYP2D6*1xN*, occur at a frequency of 2 – 12%.[92] If *CYP2D6* pharmacogenetic tests do not assess structural variation, a relatively large number of people may be assigned an incorrect diplotype.

The goal of this study was to compare the predictive power of *CYP2D6* diplotype and activity score assignments based on single nucleotide variation (SNV) and small insertion or deletion (indel) data alone and assignments using Stargazer, a software tool for assigning alleles from next-generation sequencing data that combines SNV/indel calls with structural variation data to improve the predictive model. We assessed *CYP2D6* activity in a panel of 309 human livers using CYP2D6 probe substrates, metoprolol and dextromethorphan. Metoprolol is metabolized to α-hydroxymetoprolol and

dextromethorphan is metabolized to dextrorphan via CYP2D6. We then used PGRNSeq, a next generation sequencing platform, to assign *CYP2D6* diplotypes; we compared assignments from SNV data alone and incorporation of structural variation identified from Stargazer. Finally, we compared the amount of variation in CYP2D6 activity predicted by activity score assignments from the two methods.

2.2 Materials and methods

2.2.1 Chemicals and reagents

Dextromethorphan, metoprolol tartrate, NADPH, potassium phosphate, EDTA, and acetonitrile were purchased from Sigma-Aldrich (St Louis, MO, USA). Dextrorphan-d3 was purchased from Cerilliant (Round Rock, TX, USA). Carvedilol-d3 was purchased from C/D/N Isotopes (Pointe-Claire, QC, CA).

2.2.2 Human liver samples

2.2.2.1 Human liver microsomes, total protein quantitation, cytochrome P450 oxidoreductase protein quantitation, genomic DNA extraction

Human liver samples (n = 347) were obtained from the University of Washington Human Liver Bank (n = 65) and the St. Jude Liver Resource at St. Jude Children's Research Hospital (n = 282) as described previously. HLM preparation and total protein quantitation were previously described.[98] A subset of 309 HLMs were used in this study, with 47 from the University of Washington and 262 from St. Jude liver banks.

2.2.2.2 CYP2D6 and POR protein quantitation and CYP2D6, POR, and AKR1D1 mRNA quantitation

Microsomal CYP2D6 and POR protein content was quantitated using a surrogate peptide-based LC-MS/MS method that has been previously described. [98,99] Methods for

RNA isolation, TruSeq stranded mRNA preparation, read processing, and analysis have also been previously described.[100]

2.2.2.3 DNA isolation, CYP2D6 sequencing, and CYP2D6 allele and activity score assignment

The PGRNSeq platform was used to identify *CYP2D6* genomic variation. All *CYP2D6* exons as well as the regions 2 kb upstream and 1 kb downstream were sequenced using the PGRNseq platform, as previously described.[93] Allele and subsequent star-allele diplotype assignments were first made manually using SNV and indel data only and based on criteria from the Pharmacogene Variation Consortium (https://www.pharmvar.org/gene/CYP2D6). The Stargazer algorithm (paper submitted) was then used to detect structural variation in the *CYP2D6* gene in PGRNseq sequencing data and the star-allele diplotype assignments determined manually were corrected accordingly. Activity scores were assigned to each allele based on CPIC criteria.[101]

2.2.2.4 Measurement of CYP2D6 metabolite formation rate in HLMs

HLM incubations were conducted using a dextromethorphan concentration of 1.5 μM and a metoprolol concentration of 4 μM. Optimization experiments confirmed that these concentrations were in the linear range of metabolite formation; additional optimization conditions confirmed that incubation time and total HLM protein content were also within linear conditions (data not shown). Incubations were performed in triplicate using 20 μg of total HLM protein per well diluted in pH 7.4 100 mM potassium phosphate and 1 mM EDTA buffer. HLMs were pre-incubated with substrate for 5 min at 37°C. NADPH (final concentration 1 mM) was added to start the reaction and samples

were incubated for 20 min (metoprolol) and 30 min (dextromethorphan) at 37°C. Reactions were terminated by the addition of ice-cold acetonitrile.

2.2.2.5 Metabolite and parent drug quantitation

For metoprolol and α-hydroxymetoprolol, calibration standards were prepared over the range of 0.003 to 0.176 μM; carvedilol-d4 (1 ng/μl) was used as an internal standard. Ions m/z 284.1 and 411.1 were monitored for α-hydroxyl-metoprolol and carvedilol-d4, respectively. The mobile phase was 10mM ammonium formate, pH 4, in water (A) and acetonitrile (B) using the following gradient: 0 min 15% B, 3.5 min 60%B, 4.5 min 90%B, 5 min 90%B, 5.1 min 15%B, and 9.5 min 15%B.

For dextromethorphan and dextrorphan, calibration standards were prepared over the range of 0.004 to 1 μM; dextrorphan-d3 (0.5 ng/μl) was used as an internal standard. Ions m/z 258.1 and 261.1 were monitored for dextrorphan and dextrorphan-d3, respectively. The mobile phase was 0.1% formic acid in water (A) and acetonitrile (B) using the following gradient: 0 min 20% B, 1 min 20% B, 5 min 40% B, 6 min 20% B, and 9 min 20% B.

Quantitation of dextromethorphan dextrorphan, metoprolol, and α-hydroxymetoprolol was performed on an Agilent Technologies G1956B mass spectrometer coupled to an Agilent 1200 series HPLC using a Zorbax SB-C18 2.1 mm x 150 mm x 5 μm column (Agilent Technologies, Santa Clara, CA). For dextrorphan, the column was maintained at 30°C with a flow rate of 0.3 ml/min. For α-hydroxymetoprolol, the column was maintained at 35°C with a flow rate of 0.25 ml/min.

2.2.2.6 Statistical analysis

Samples were included for statistical analysis if both PGRNSeq and activity data for at least one probe substrate were available (n =309). All analyses were performed using RStudio 1.0.143 (Boston, MA). In all analyses, liver collection site was included as a covariate due to significant differences in mean enzyme activity between the sites as well as differences in POR mRNA and protein content and AKR1D1 mRNA content. Pairwise associations between donor demographics (e.g. sex, ethnicity, age), liver microsomal CYP2D6 and POR protein content, liver microsomal *CYP2D6, POR,* and *AKR1D1* mRNA content, and liver microsomal CYP2D6 activity (dextrorphan and α-hydroxymetoprolol formation rate) were tested with linear regression using robust standard errors. This approach was also used to test the association between activity score and CYP2D6 activity. A multivariate linear regression model with robust standard errors was used to assess the contribution of selected predictors to the variability observed in CYP2D6 activity.

2.3 Results

2.3.1 Liver donor demographics

Liver donor demographics (n = 309) are presented in Table 1. Donors ranged from 0 to 87 years old, with an average age of 39.6 years. A majority of the donors were male (58.3%) and almost all were of European descent (95.5%). Liver tissue was collected from donors at two sites: the University of Washington in Seattle, WA (n = 47) and St. Jude Children's Research Hospital in Memphis, TN (n = 262). Data on liver pathology and/or disease were not collected for all donors; similarly, medication lists are missing for some donors and possibly incomplete for others.

2.3.2 Effects of CYP2D6 mRNA and protein content on activity

The rates of α-hydroxymetoprolol and dextrorphan formation, used as a measure of CYP2D6 activity, were both significantly correlated with *CYP2D6* mRNA content (Figures 3A and 3B) and CYP2D6 protein content (Figures 4A and 4B). For association between metabolite formation rates and *CYP2D6* mRNA content, R^2 values were 0.231 and 0.254 for α-hydroxymetoprolol and dextrorphan formation rates, respectively. CYP2D6 protein content had a much stronger association with CYP2D6 activity, as reflected in higher R^2 values of 0.643for α-hydroxymetoprolol formation rate and 0.751for dextrorphan formation rate. It should be noted that protein data were only available for 98 of the 309 samples analyzed, 10 of which were below the lower level of quantitation. With collection site included as a covariate, *CYP2D6* mRNA content was not a significant predictor of CYP2D6 protein content.

2.3.3 CYP2D6 allele and diplotype frequencies in the liver bank samples

We identified 23 alleles in the liver bank samples using the Stargzer algorithm (Table 2) and 69 unique diplotypes (Table 3). Allele frequencies reflect the predominantly European descent of the liver bank samples. The three most common alleles were *CYP2D6*1* (32%), *CYP2D6*2* (14.7%), and *CYP2D6*4* (13.8%), which is consistent with published reports for populations of European descent.[85]

To summarize, a total of seven different alleles, including *CYP2D6*1* and *CYP2D6*2,* were identified that confer normal enzyme function and a normal metabolizer (NM) phenotype (activity score of 1). The seven NM alleles occurred at an aggregate frequency of 54.2%. Six alleles were identified with reduced enzyme function and an intermediate metabolizer (IM) phenotype (activity score of 0.5). The most

common IM allele was *CYP2D6*41* (11.2%), a splice variant allele found most commonly in European populations. The IM alleles occurred at an aggregate frequency of 18.9%. Seven loss of function alleles were identified conferring a poor metabolizer (PM) phenotype (activity score of 0), occurring at an aggregate frequency of 25.7%. The most common PM allele was *CYP2D6*4* (13.8%). Two alleles associated with the ultrarapid metabolizer (UM) phenotype (activity score of 2) were identified, *CYP2D6*1x2* (0.5%) and *CYP2D6*2x2* (0.6%).

2.3.4 Structural variation data is critical for correct diplotype assignment

A total of 16 unique alleles were identified from SNV/indel data alone. The inclusion of data on structural variation allowed for the identification of 7 additional alleles with gene deletions, gene duplications, and hybrid tandem gene arrangements, bringing the total to 23 unique alleles (Table 2).

Activity scores were assigned from diplotypes identified from SNV/indel data alone (16 alleles) and from diplotypes identified with the inclusion of structural variation data (23 alleles). When the samples were analyzed incorporating structural variation data, 67 samples (22%) were incorrectly assigned with SNV/indel data alone. Importantly, inclusion of structural variation data changed the predicted activity score for 26 of those samples, representing approximately 8% of the liver bank donors. Figure 1 describes the diplotypes of the 67 samples with structural variation alleles that were identified with Stargazer as well as changes in activity score resulting from changes in predicted activity score.

*CYP2D6*5*, a complete gene deletion, was identified in 23 samples, occurring at an allele frequency of 3.7%. The importance of including the *5 allele is evident as the

corrected diplotype calls decreased the activity score of 18 samples in this sample set (5.8%). The remaining five samples already had activity scores of 0 and thus could not be decreased further.

A total of 8 duplicated alleles in the form of *CYP2D6*xN* were identified, consisting of *CYP2D6*1x2*, *CYP2D6*2x2*, and *CYP2D6*41x2*. As these are all duplications of functional alleles, they resulted in activity score increases for all of the samples they were identified in.

Structural data were also used to detect the hybrid tandem alleles *CYP2D6*4N+*4, CYP2D6*68+*4*, and CYP2D6*36+*10*. Inclusion of hybrid tandem alleles did not result in any activity score changes in our sample set. *CYP2D6*68+*4* is a non-functional allele identified in samples previously designated as **4* with SNV/indel data alone, and were already assigned an activity score of 0. *CYP2D6*36+*10* was identified in samples previously designated **10*, which both have an activity score of 0.5. Published data suggest that there is no significant difference in enzyme activity conferred by *CYP2D6*36+*10* and **10* alone[102]; thus, the activity score did not change from 0.5.

The inclusion of structural data in *CYP2D6* significantly improved the predictive value of the activity score assignments for both probe substrates, metoprolol and dextromethorphan (Figure 2; Table 4). Using SNV/indel data alone, activity score predicted 31.5% and 36.4% of the variability in CYP2D6 activity when using metoprolol and dextromethorphan as probes, respectively (Table 4A). With structural variation data included from Stargazer, these numbers increased to 36.6% and 40.7% (Table 4B).

2.3.5 Liver collection site, POR protein content, and *AKR1D1* mRNA content contribute to variability in CYP2D6 activity

Liver collection site was highly predictive of CYP2D6 activity, consistent with our collaborators reports from these liver bank samples; therefore, collection site was included in all analyses as a covariate. [98,100] POR protein content, but not mRNA content, was significantly associated with metabolite formation rate for both α-hydroxymetoprolol ($R^2 = 0.198$, $p < 0.001$) and dextrorphan formation rate ($R^2 = 0.219$, $p < 0.001$) (Figure 5). *AKR1D1* mRNA content was also associated with CYP2D6 activity for both α-hydroxymetoprolol ($R^2 = 0.16$, $p < 0.001$) and dextrorphan formation rate ($R^2 = 0.208$, $p < 0.001$) (Figure 6).

2.3.6 Multivariate regression analysis of CYP2D6 activity

A multivariate linear regression using robust standard errors was performed to determine which variables, in addition to activity score, explain some of the variability in *CYP2D6* activity (Table 5). Donor age and sex were not included in the model due to a lack of significance in primary analysis, which is consistent with the literature.[88] While ethnicity is known to make a large contribution to variation in CYP2D6 activity, we did not include it as a covariate in our model because the liver donors were almost entirely of European descent. We also excluded liver disease and concomitant CYP2D6 inhibitors because the data were incomplete for the majority of donors.

The R^2 for α-hydroxymetoprolol formation rate was 0.471, an improvement over the value of 0.364 obtained when just activity score and collection site were used as predictors. The R^2 for dextrorphan formation rate was 0.505, also an improvement over the value of 0.407 obtained when just activity score and collection site were used as predictors. The covariates that contributed to these increases in R^2 were POR protein

content and *AKR1D1* mRNA content. Similar to what was observed in the model that included just activity score and collection site, activity scores of 2 and 3 were not significant predictors of CYP2D6 activity.

2.4 Discussion

We found that inclusion of structural variation data is essential to correctly assigning *CYP2D6* activity scores and these data are more predictive of CYP2D6 activity than activity scores assigned based on SNV/indel data alone. Accurate identification of gene deletions, gene duplications, and hybrid alleles will maximize the clinical utility of *CYP2D6* pharmacogenetic tests. We also found that POR protein content is correlated with CYP2D6 activity. As *POR* is a polymorphic gene, inclusion of *POR* pharmacogenetic data in the clinical setting may further improve *CYP2D6* phenotype prediction.

When diplotypes were assigned based on SNV/indel data alone, resultant activity scores explained about 31% of the variation in CYP2D6 activity when metoprolol was used as a probe and 36% of the variation when dextromethorphan was used. When diplotypes were reassigned based on the Stargazer algorithm that utilizes both SNV/indel and structural variation data, this number increased to approximately 36% for metoprolol and 41% for dextromethorphan. The diplotypes of 67 samples were reassigned, representing over 20% of the samples tested. Importantly, the activity score changed for 26 samples, representing approximately 8% of the samples tested. The samples that had incorrect activity scores contained alleles that contribute to “extreme” phenotypes: the whole-gene deletion *CYP2D6*5* or duplications of functional alleles *CYP2D6*1*, or *2. The hybrid alleles **4N*, **36*, and **68* are nonfunctional and thus did not affect activity

score assignments when in tandem arrangements with alleles identified via SNV/indel analysis (**4* or **10*). These tandem hybrid alleles are still important to identify, though, since their presence is informative when resolving potentially ambiguous diplotypes. For example, many commercially available tests would detect a duplication in an individual with a **1/*4N + *4* diplotype but would be unable to determine if the duplication is on a **1* or **4* allele.

Due to the lack of ethnic diversity in the liver bank donors, we were unable to assess for differences in *CYP2D6* activity that may be attributable to ethnicity. Though very few liver donors were of non-European descent, however, their presence is reflected in some of the more uncommon alleles. The **17* allele occurs at a frequency <1% in European populations and an average of 18.2% in African Americans.[85] In this study, the allele was present in four donors, two of them of known African descent. An African-American donor also contributed the single **29* allele, which occurs at frequencies of <1% in those of European descent and 6.5% in those of African descent.[85] In East Asian populations, the **10* allele is very common, occurring at an average frequency of 42.6%. The *CYP2D6*10* allele is also found in tandem with the hybrid *CYP2D6*36* allele as *CYP2D6*36+*10*, found most frequently in East Asian populations at an average frequency of 26.4%.[85,103] These high frequency alleles were observed in the *CYP2D6*10/*36+*10* diplotype of the only Asian donor included in this study.

CYP2D6 activity was significantly correlated with both *CYP2D6* mRNA and protein contents. The correlation with *CYP2D6* mRNA was weaker, suggesting that mRNA content does not explain very much of the variability in CYP2D6 activity. The correlation with protein, on the other hand, was much stronger, which is unsurprising

since the CYP2D6 enzyme content is more proximal to metabolite formation rate than mRNA in the regulatory scheme. The correlation between *CYP2D6* mRNA and protein was not significant, which fits a general, genome wide trend of low mRNA-protein correlations.[104]

The correlation between CYP2D6 activity and POR protein content was significant for both substrates. Further investigation into POR diplotype for each sample would be worthwhile. Because POR is required for CYP enzyme function, genetic variation in the *POR* gene may explain some of the variability in CYP activity not explained by CYP genetics. Previous data supports a possible link between POR and CYP2D6. A study examining the effects of various recombinant *POR* variants on CYP2D6 activity found that the common *POR* variant A503V decreased CYP2D6 activity by 40-50% when using dextromethorphan and bufurolol as probes.[90] Thus, *POR* genotyping may provide additional information that can be used with *CYP2D6* diplotype data in a clinical setting to improve CYP2D6 phenotype predictions.

AKR1D1 mRNA content was very weakly correlated with CYP2D6 activity. AKR1D1 protein content data were not available for comparison. AKR1D1 is involved in bile acid homeostasis and is hypothesized to regulate CYPs through activation of nuclear receptors that transcriptionally regulate CYP expression. Bile acids are ligands of nuclear receptors PXR and CAR, which induce the transcription of *CYP3A4*, *CYP2C9,* and *CYP2C19.* Since *CYP2D6* is generally considered to be non-inducible, this mechanism of CYP regulation should not affect CYP2D6 activity. It is possible, however, that there are additional regulatory pathways that have not yet been described.[91] Interestingly, *AKR1D1* mRNA remained a significant predictor for both substrates, suggesting that AKR1D1

might affect CYP2D6 through the hypothesized mechanism or through a mechanism that has not yet been described.

The liver procurement site explained a relatively large amount of variability in CYP2D6 activity (7.8% for metoprolol, 12.8% for dextromethorphan) and was a significant predictor for both probe substrates. Other groups have noted similar differences when using the liver bank to study *CYP2A6* and *CYP2C19*. [98,100] Factors such as liver tissue collection method and storage conditions are some of the many variables that may have contributed to the site difference.

A multivariate linear regression model was used to assess the combined contributions of activity score, POR protein content, *AKR1D1* mRNA content, and liver procurement site to variability in CYP2D6 activity. All of the covariates were significant predictors of CYP2D6 activity except for activity scores 2 and 3. Compared to the other activity score categories, there was a higher degree of variability in CYP2D6 activity for activity scores 2 and 3. Also, the metabolite formation rates for samples with an activity score of 3 were lower than expected and we are unable to explain this with the PGRNSeq data. Some of this may have been due to the small number of samples (n = 3) with an activity score of 3. These findings could also be due to genetic variation that has yet to be described, either in the *CYP2D6* gene (although that is less likely with the methodology employed), regulatory regions, genes related to CYP2D6 function such as *POR*, or due to co-administration of CYP2D6 inhibitors not recorded in donor medication lists.

Together, data on CYP2D6 activity score, POR protein content, *AKR1D1* mRNA content, and liver procurement site explained 47% and 51% of the variability in CYP2D6 activity when metoprolol and dextromethorphan were used as probes, respectively.

Limitations of this study include small sample sizes in the activity score 2.5 (n=2) and 3 (n=3) categories and a relatively large amount of variation in the activity score 2 and 3 categories. One sample in particular, genotyped as a **1/*1x2* and assigned an activity score of 3, had very low *CYP2D6* activity, consistent for both probe drugs. It's unclear if this is due to unidentified genetic variation in *CYP2D6*, post-transcriptional regulation, unrecorded use of drugs that inhibit CYP2D6, or issues such as low HLM quality.

Most currently available *CYP2D6* pharmacogenetic tests are designed to detect a limited selection of alleles with SNVs and indels, the whole-gene deletion *CYP2D6*5,* and nonspecific duplication signals (**xN).* Our approach of using Stargazer to analyze targeted exome sequencing data offers a solution to many of the problems with current tests. First, we can detect uncommon and rare alleles not included on targeted genotyping panels, solving the problem of defaulting to a call of *CYP2D6*1* or *CYP2D6*2* when a functional allele is present but not tested for. Second, our approach allows for identification of specific duplicated alleles rather than a potentially ambiguous nonspecific duplication signal. Third, the ability to identify hybrid genes in tandem arrangements with *CYP2D6* solves the problem of incorrect duplication calls when using real-time PCR *CYP2D6* copy number assays. Furthermore, identification of hybrid alleles that do not occur in tandem with *CYP2D6* is an improvement over PCR-based genotyping methods that may over-amplify the non-hybrid allele, making the sample appear to be homozygous for the non-hybrid.[75]

For a gene as complicated as *CYP2D6*, analysis of SNVs and indels alone is not sufficient for making accurate allele calls. If we are to strive for precision, structural

variation in the *CYP2D6* gene must also be analyzed beyond *CYP2D6*5* and nonspecific duplication signals.

2.5 Tables

2.5.1 Table 1. Liver donor demographics (n = 309)

Age		
Mean ± SD		39.6 ± 22.4
Range		0 - 87
Sex		
Male		180 (58.3%)
Female		124 (40.1%)
Unknown		5 (1.6%)
Ethnic background		
European		295 (95.5%)
African		9 (2.9%)
Asian		1 (0.3%)
Hispanic		1 (0.3%)
Unknown		3 (1.0%)
Liver disease		
No / not recorded		256 (82.8%)
Yes		53 (17.2%)
	Fatty liver	25 (47.2%)
	Cancer	7 (13.2%)
	Fibrosis	5 (9.4%)
	Acute injury	4 (7.5%)
	Cirrhosis	4 (7.5%)
	Hepatitis	4 (7.5%)
	Biliary atresia	4 (7.5%)
CYP2D6 inhibitor		
No / not recorded		295 (95.5%)
Yes		14 (26.4%)
	Cimetidine	12 (85.7%)
	Amiodarone	1 (7.1%)
	Fluoxetine	1 (7.1%)

2.5.2 Table 2. *CYP2D6* alleles (n = 618) identified in liver bank samples

Allele	n (%)	Nucleotide Changes	Effect	Phenotype		Activity Score
				Activity	Class†	
Single nucleotide variants and indels						
***1**	198 (32%)	N/A	N/A	Normal	NM	1
***2**	91 (14.7%)	2850C>T, 4180G>C	R296C; S486T	Normal	NM	1
***33**	5 (0.8%)	2483G>T	A237S	Normal	NM	1
***35**	37 (6%)	31G>A, 2850C>T, 4180G>C	V11M; R296C; S486T	Normal	NM	1
***39**	1 (0.2%)	1661G>C, 4180G>C	S486T	Normal	NM	1
***43**	2 (0.3%)	77G>A	R26H	Normal	NM	1
***9**	21 (3.4%)	2615_2617delAAG	K281del	Decreased	IM	0.5
***10**	15 (2.4%)	100C>T, 4180G>C	P34T; S486T	Decreased	IM	0.5
***17**	4 (0.6%)	1023C>T, 2850C>T, 4180G>C	T197I; R296C; S486T	Decreased	IM	0.5
***29**	1 (0.2%)	1659G>A, 2850C>T, 3183G>A, 4180G>C	V136I; R296C; V338M; S486T	Decreased	IM	0.5
***41**	69 (11.2%)	2850C>T, 2988G>A, 4180G>C	Splicing defect	Decreased	IM	0.5
***59**	5 (0.8%)	2291G>A, 2850C>T, 2939G>A, 4180G>C	Decreased mRNA expression; R296C; S486T	Decreased	IM	0.5
***3**	7 (1.1%)	2549delA	Frameshift	None	PM	0
***4**	85 (13.8%)	100C>T, 1846G>A, 4180G>C	Splicing defect	None	PM	0
***6**	4 (0.6%)	1707delT	Splicing defect	None	PM	0
***20**	3 (0.5%)	1973_1974insG, 1978C>T, 1979T>C, 2850C>T, 4180G>C	Frameshift	None	PM	0
Structural variants						
***1x2**	3 (0.5%)	Duplication of *CYP2D6*1*	Two active genes	Increased	UM	2
***2x2**	4 (0.6%)	Duplication of *CYP2D6*2*	Two active genes	Increased	UM	2
***41x2**	1 (0.2%)	Duplication of *CYP2D6*41*	Two decreased activity genes	Normal	NM	1
***36+*10**	2 (0.3%)	*10 in tandem with *CYP2D7/CYP2D6* hybrid gene	One decreased activity and one inactive gene	Decreased	IM	0.5
***4N+*4**	4 (0.6%)	*4 in tandem with *CYP2D7/CYP2D6* hybrid gene	Two inactive genes	None	PM	0
***68+*4**	33 (5.3%)	*4 in tandem with *CYP2D7/CYP2D6* hybrid gene	Two inactive genes	None	PM	0
***5**	23 (3.7%)	*CYP2D6* deleted	*CYP2D6* deleted	None	PM	0

† UM: ultrarapid metabolizer, NM: normal metabolizer, IM: intermediate metabolizer, PM: poor metabolizer

2.5.3 Table 3. *CYP2D6* diplotypes (n = 309) identified in liver bank samples

Diplotype	n	Activity score	Diplotype	n	Activity score
1 / 1x2	1	3	1 / 5	8	1
2 / 1x2	1	3	1 / 6	1	1
35 / 2x2	1	3	1 / 20	1	1
41 / 2x2	2	2.5	2 / 3	1	1
1 / 33	2	2	2 / 4	13	1
2 / 33	1	2	2 / 5	4	1
1 / 35	12	2	2 / 6	1	1
2 / 35	2	2	9 / 9	1	1
1 / 1	29	2	1 / 4 + 68	12	1
1 / 2	35	2	1 / 4N + 4	2	1
2 / 2	5	2	10 / 10 + 36	1	1
35 / 35	1	2	2 / 4 + 68	5	1
4 / 1x2	1	2	20 / 35	1	1
4 / 2x2	1	2	35 / 4 + 68	2	1
10 / 35	4	1.5	41 / 41	1	1
1 / 41	18	1.5	41 / 59	1	1
2 / 41	14	1.5	4N + 4 / 33	1	1
1 / 59	1	1.5	4 / 41	10	0.5
2 / 59	3	1.5	5 / 41	3	0.5
1 / 9	8	1.5	6 / 41	1	0.5
1 / 10	8	1.5	4 / 9	3	0.5
1 / 17	3	1.5	4 / 10	1	0.5
2 / 9	1	1.5	4 / 17	1	0.5
10 / 41x2	1	1.5	4 + 68 / 10 + 36	1	0.5
29 / 39	1	1.5	41 / 4 + 68	3	0.5
35 / 41	8	1.5	41 / 4N + 4	1	0.5
4 / 33	1	1	9 / 4 + 68	1	0.5
4 / 35	3	1	3 / 5	1	0
5 / 35	1	1	4 / 4	7	0
6 / 35	1	1	4 / 5	4	0
9 / 41	6	1	4 / 20	1	0
3 / 43	1	1	3 / 4 + 68	2	0
4 / 43	1	1	4 / 4 + 68	5	0
1 / 3	2	1	5 / 4 + 68	2	0
1 / 4	26	1			

2.5.4 Table 4. Multiple linear regression: comparison of association between CYP2D6 activity and activity score when activity score is assigned using SNV/indel data only (A) and SNV/indel plus structural variation data (B)

A. Prediction of CYP2D6 activity using activity scores assigned based on SNV/indel data only

	Metoprolol			Dextromethorphan		
	β	*SE*	*p*	*β*	*SE*	*p*
Intercept	10.513	1.340	**< 0.001**	61.056	6.160	**< 0.001**
AS 0	-10.778	1.230	**< 0.001**	-65.220	6.040	**< 0.001**
AS 0.5	-9.582	1.290	**< 0.001**	-58.369	6.000	**< 0.001**
AS 1	-6.234	1.160	**< 0.001**	-42.609	5.480	**< 0.001**
AS 1.5	-3.420	1.370	**0.013**	-32.548	6.220	**< 0.001**
AS 2	0.977	1.690	0.564	-11.587	7.540	0.1253
Site: UW	5.252	1.340	**< 0.001**	30.554	6.160	**< 0.001**
Degrees of freedom	285			302		
R^2 / adj. R^2	0.3149 / 0.3005			0.3643 / 0.3516		

B. Prediction of CYP2D6 activity using activity scores assigned with Stargazer, based on both SNV/indel and structural variation data

	Metoprolol			Dextromethorphan		
	β	*SE*	*p*	*β*	*SE*	*p*
Intercept	10.917	1.280	**< 0.001**	62.506	5.840	**< 0.001**
AS 0	-11.074	1.180	**< 0.001**	-66.055	5.720	**< 0.001**
AS 0.5	-9.957	1.230	**< 0.001**	-59.781	5.920	**< 0.001**
AS 1	-6.154	1.150	**< 0.001**	-42.283	5.310	**< 0.001**
AS 1.5	-4.357	1.260	**< 0.001**	-36.042	5.680	**< 0.001**
AS 2	1.298	1.720	0.450	-9.978	7.580	0.190
AS 2.5	12.840	1.210	**< 0.001**	30.092	8.060	**< 0.001**
AS 3	-2.735	4.110	0.506	-15.124	26.430	0.570
Site: UW	4.848	1.280	**< 0.001**	29.105	5.840	**< 0.001**
Degrees of freedom	283			300		
R^2 / adj. R^2	0.3636 / 0.3456			0.407 / 0.3912		

2.5.5 Table 5. Multiple linear regression: association between CYP2D6 activity and activity score, POR protein content, and *AKR1D1* mRNA content

	Metoprolol			Dextromethorphan		
	β	*SE*	*p*	*β*	*SE*	*p*
Intercept	4.540	1.792	**0.012**	38.133	7.316	**< 0.001**
AS 0	-11.162	1.452	**< 0.001**	-68.953	7.112	**< 0.001**
AS 0.5	-9.147	1.393	**< 0.001**	-58.065	6.508	**< 0.001**
AS 1	-5.277	1.090	**< 0.001**	-39.621	5.240	**< 0.001**
AS 1.5	-3.993	1.299	**0.002**	-37.050	5.697	**< 0.001**
AS 2	1.660	1.727	0.338	-9.388	7.694	0.224
AS 2.5	12.219	0.593	**< 0.001**	41.358	2.468	**< 0.001**
AS 3	-0.569	2.898	0.845	-7.004	22.916	0.760
Site: UW	4.396	1.221	**< 0.001**	27.737	5.739	**< 0.001**
POR protein	0.185	0.034	**< 0.001**	0.697	0.136	**< 0.001**
AKR1D1 mRNA	0.038	0.012	**0.002**	0.218	0.053	**< 0.001**
Degrees of freedom	223			237		
R^2 / adj. R^2	0.4712 / 0.4474			0.505 / 0.4941		

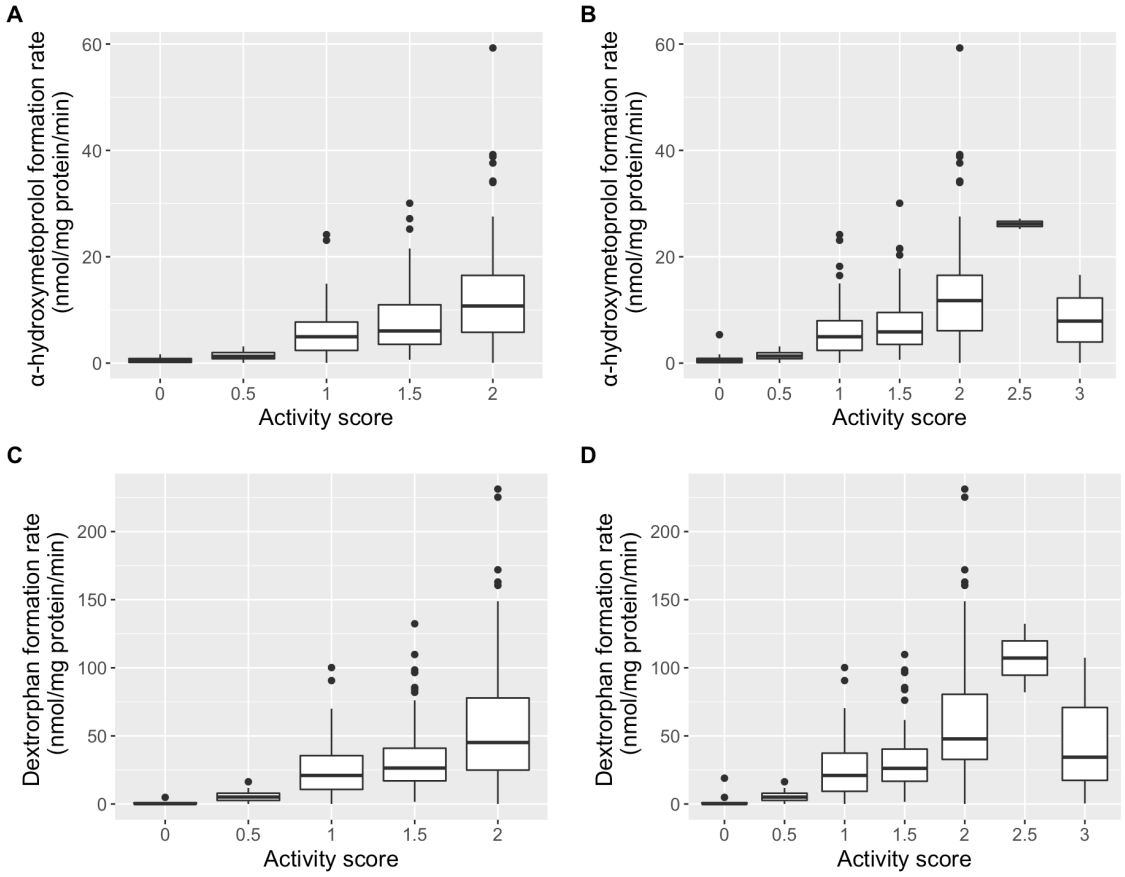
2.6 Figures

2.6.1 Figure 1. Comparison of diplotype and activity score changes assigned with SNV/indel data alone and with inclusion of structural variation data with Stargazer.

SNV data alone			SNV + structural variation data			
Diplotype	**AS**	**n**	**Diplotype**	**AS**		**n**
Deletion						
*1/*1	2	38	*1/*5	ê	1	8
*2/*2	2	9	*2/*5	ê	1	4
*35/*35	2	2	*5/*35	ê	1	1
*41/*41	1	4	*5/*41	ê	0.5	3
*4/*10	0.5	4	*4/*5	ê	0	2
*3/*3	0	1	*3/*5	è	0	1
*4/*4	0	16	*4/*5	è	0	2
Duplication						
*1/*1	2	38	*1/*1x2	é	3	1
*1/*2	2	36	*1x2/*2	é	3	1
*2/*35	2	3	*2x2/*35	é	3	1
*2/*41	1.5	16	*2x2/*41	é	2.5	2
*1/*4	1	41	*1x2/*4	é	2	1
*2/*4	1	19	*2x2/*4	é	2	1
*10/*41	1	1	*10/*41x2	é	1.5	1
Hybrid tandem						
*1/*4	1	41	*1/*68 + *4	è	1	12
			*1/*4N + *4	è	1	2
*2/*4	1	19	*2/*68 + *4	è	1	5
*4/*35	1	5	*35/*68 + *4	è	1	2
*4/*33	1	2	*4N + 4/*33	è	1	1
*10/*10	1	1	*10/*36 + *10	è	1	1
*4/*9	0.5	4	*9/*68 + *4	è	0.5	1
*4/*10	0.5	4	*36 + *10/*68 + *4	è	0.5	1
*4/*41	0.5	14	*41/*68 + *4	è	0.5	3
			*4N + *4/41	è	0.5	1
*3/*4	0	2	*3/*68 + *4	è	0	2
*4/*4	0	16	*4/*68 + *4	è	0	5
Deletion + hybrid tandem						
*4/*4	0	16	*5/*68 + *4	è	0	2

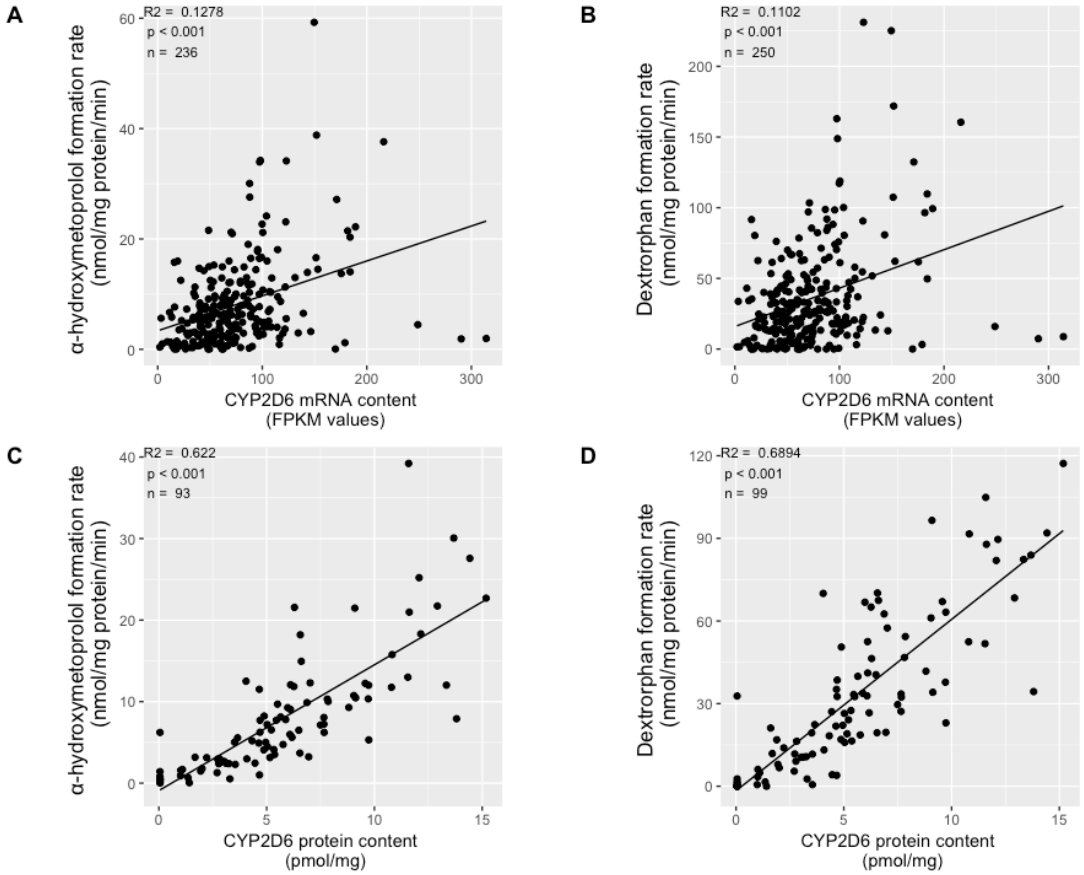
Columns on the left show selected diplotypes and activity scores (AS) assigned using allele calls from SNV/indel-only data. Corrected diplotypes and activity scores, based on Stargazer allele assignments, are displayed in the columns on the right. Activity scores are color-coded: 3 and 2.5 (dark green), 2 (green), 1 (yellow), 0.5 (orange), and 0 (red). Arrows indicate the direction of the change in activity score assignment with the incorporation of structural data: decrease (↓), increase (↑), and no change (→).

2.6.2 Figure 2. Association between CYP2D6 metabolite formation rate and CYP2D6 activity score.



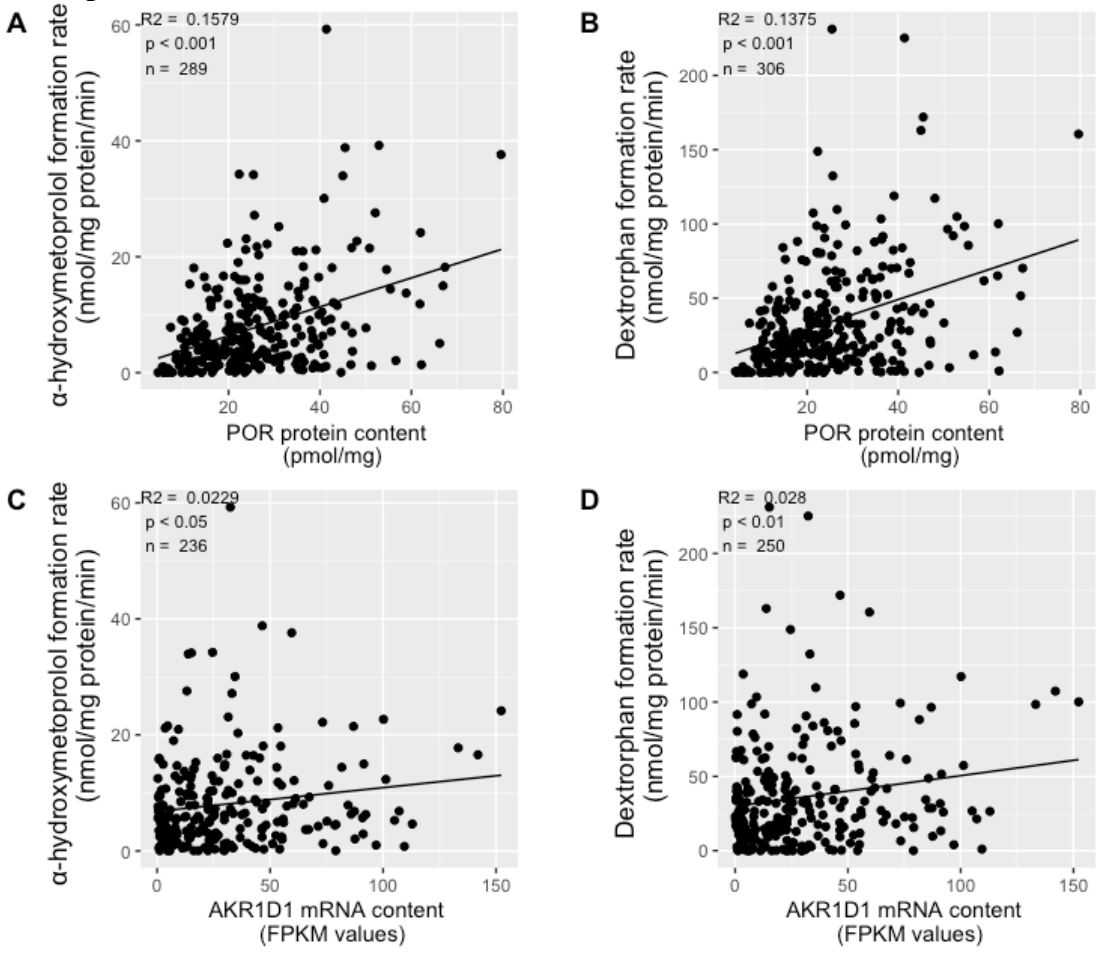
Panels A and B show alpha-hydroxymetoprolol formation rate by activity score assigned with SNV/indel data alone (panel A) and with Stargazer (panel B). Panels C and D show dextrorphan formation rate formation rate but activity score assigned with SNV/indel data alone (panel C) and with Stargazer (panel D).

2.6.3 Figure 3. Association between CYP2D6 metabolite formation rate and *CYP2D6* mRNA and protein content.



CYP2D6 metabolite formation rate was associated with *CYP2D6* mRNA content quantitated by RNASeq: alpha-hydroxymetoprolol formation rate (panel A) and dextrorphan formation rate (panel B); CYP2D6 metabolite formation rate was associated with CYP2D6 protein content quantitated by LC-MS/MS: alpha-hydroxymetoprolol formation rate (panel C) and dextrorphan formation rate (panel D).

2.6.4 Figure 4. Association between CYP2D6 metabolite formation rate and POR protein and *AKR1D1* mRNA content.



CYP2D6 metabolite formation rate was associated with POR protein content quantitated by LC-MS/MS: alpha-hydroxymetoprolol formation rate (panel A) and dextrorphan formation rate (panel B); CYP2D6 metabolite formation rate was associated with *AKR1D1* mRNA content quantitated by RNASeq: alpha-hydroxymetoprolol formation rate (panel C) and dextrorphan formation rate (panel D).

3 Discussion

3.1 Project summary

The goal of my research was to identify rare coding and structural variation that affects CYP2D6 activity that is not yet routinely included in pharmacogenetic tests used in the clinic and demonstrate how the variants can improve the predictive ability of CYP2D6 activity score. To accomplish this goal, I used human liver microsomes and DNA sequencing data derived from a human liver bank to determine the correlation between activity score and *CYP2D6* before and after incorporating structural and copy number variation data derived from the Stargazer platform.

In Specific Aim 1, I used data from the PGRNseq platform to assign diplotypes to 309 human liver bank donors. Part of this aim was to identify rare variants of functional consequence. While I did identify uncommon (<5% allele frequency) and rare (<1% allele frequency) variants, none of them appeared to cause a dramatic change in activity. The alleles that I identified in the liver bank samples occurred at frequencies that reflected the mostly-European demographics of the donors.

In Specific Aim 2, I used human liver microsomes (HLMs) as an *in vitro* model to determine CYP2D6 activity. For probe drugs, I chose to use dextromethorphan because it is well established as a probe in the CYP2D6 literature and metoprolol because of its clinical relevance. Using the metabolite formation rates from the HLM incubations, I found that activity score explained 31% (using metoprolol) to 36% (using dextromethorphan) of the variability in CYP2D6 activity in the liver bank samples when diplotypes were assigned using SNV and indel data alone.

Specific Aim 3 was added after Specific Aims 1 and 2 were complete. I used structural and copy number variation data from the Stargazer algorithm to assign new diplotypes to samples that could not be identified as having structural variations based on SNVs and indels alone. Approximately 20% of the samples were assigned new diplotypes and 8% were assigned a new activity score as a result.

After seeing these changes in activity score assignments, the next step was to assess whether activity score was actually more predictive of CYP2D6 function. When diplotypes were reassigned based on the Stargazer algorithm that utilizes both SNV/indel and structural variation data, this amount of variability in CYP2D6 activity explained by activity score increased to 36% for metoprolol and 41% for dextromethorphan.

I then created a multiple linear regression model using the activity data generated in Specific Aim 2 and activity scores given to the Stargazer-updated diplotypes as well as additional data available for the liver samples: liver collection site, POR protein content, and *AKR1D1* mRNA content. With the additional data, the model explained around 47% (using metoprolol) to 50% (using dextromethorphan) of the observed variation in CYP2D6 activity.

There are two conclusions that stand out above the rest. The major conclusion is that the inclusion of structural and copy number variation in pharmacogenetic tests is extremely important. Incorporation of structural and copy number variation data with sequence data will help identify extreme phenotypes (e.g. poor and ultrarapid metabolizers) and will improve the predictive power of activity score. *CYP2D6*5,* the allele representing a deletion of the *CYP2D6* gene, occurs at a frequency of approximately 6% in African populations, 5% in East Asian and Oceanian populations,

and around 2-3% in European, South/Central Asian, and Middle Eastern populations. At the low end of the range, a 2% allele frequency means that 4% of the population carries the **5* allele. At the high end, an allele frequency of 6%, around 11% of the population carries **5*. This represents a large number of people that will be incorrectly genotyped if copy number variation is not included on a pharmacogenetic test. Incorrect classification of a poor metabolizer can have serious clinical ramifications. If, for example, a poor metabolizer is given a starting dose of a tricyclic antidepressant appropriate for an normal metabolizer, they are at higher risk of adverse effects such as confusion, muscle twitches, hypotension, and arrhythmias. [60,85,105] Ultrarapid metabolizer phenotypes are generally due to the presence of additional *CYP2D6* gene copies. The **1xN* allele is particularly common on Oceanian populations, occurring at a frequency of nearly 12%. Duplications are also relatively common Middle Eastern populations; **1xN* occurs at a frequency of 3.1% and **2xN* at 3.9%. This translates to over 20% of Oceanian individuals carrying a *1xN allele and around 6 and 8% of Middle Eastern individuals carrying a **1xN* and **2xN* allele, respectively. If an ultrarapid metabolizer is misclassified and, for example, is prescribed codeine, they are at higher risk of experiencing adverse effects such as oversedation or respiratory depression. [85,101]

The second, more minor conclusion is that considering *POR* pharmacogenetics along with *CYP2D6* pharmacogenetics may offer a way to improve the accuracy of CYP2D6 activity predictions. With a clinical setting in mind, the logistics would be simple since samples could be collected for *POR* and *CYP2D6* pharmacogenetic tests at the same time. Interpretation of how both *POR* and *CYP2D6* diplotypes impact CYP2D6 activity would require further study. A starting point would be to investigate the impact of *POR*

pharmacogenetics on CYP2D6 activity using the liver bank samples. A study examining the effects of various recombinant *POR* variants on CYP2D6 activity found that POR with common variant A503V decreased CYP2D6 activity by 40-50% when using dextromethorphan and bufurolol as probes.[90] The POR A503V variant, also known as *POR*28*, is common and is present at frequencies ranging from 19 – 37% across world populations.[106]

In conclusion, *CYP2D6* is a highly relevant, highly polymorphic, highly complicated gene. When using pharmacogenetic data to make predictions about CYP2D6 activity, analysis of SNVs and indels alone is not sufficient. The inclusion of structural variation data improves the precision of *CYP2D6* pharmacogenetic tests and is key to avoiding adverse effects or treatment failure in poor and ultrarapid metabolizers.

3.2 Emerging roles for pharmacists

Pharmacists can play essential roles in the clinical implementation of pharmacogenetic testing. As members of the healthcare team, they are uniquely qualified to act as translators of pharmacogenetic test results into actionable treatment recommendations.

3.2.1 Physician knowledge gaps

A major barrier to the clinical implementation of pharmacogenomics is a lack of physician knowledge. In a 2014 survey of US-based primary care physicians, cardiologists, and psychiatrists, only 12.6% reported being extremely or very familiar with pharmacogenomics. Of the physicians who had never ordered a pharmacogenetic test, common reasons were not knowing what test to order (69.7%) and being unsure about the clinical value of testing (51.9%).[107] A previous study reported similar results,

with only 13% of surveyed physicians stating that they felt well-informed about the role of pharmacogenomic testing in therapeutic decision-making. [108] A reason for this knowledge deficit is a lack of training in pharmacogenomics; only 11% of the respondents in the 2014 survey reported that they had received formal training in the topic. [107]

Physicians who are active participants in the pharmacogenomics implementation program at the Vanderbilt University School of Medicine also reported these types of knowledge deficits. Of 15 clinicians interviewed, a common theme was that it is difficult for them to keep up on current pharmacogenetics evidence and recommendations. Some also expressed difficulty with explaining the rationale behind pharmacogenetic testing to patients and their families.[109]

3.2.2 Pharmacist qualifications

Schools of pharmacy have begun filling in the pharmacogenomics knowledge gap by expanding the pharmacogenetics content in their curricula; the percentage of schools with pharmacogenetics curriculum has increased from 39% in 2005 to 89% in 2010. The Accreditation Council for Pharmacy Education (ACPE) has since made pharmacogenetics a required curricular component for all accredited pharmacy schools, listing it in their 2016 blueprint as a factor that should be emphasized in "evidence-based clinical decision making, therapeutic treatment planning, and medication therapy management strategies."[110] Professional societies have also recognized the opportunity for pharmacist involvement with clinical pharmacogenomics. For example, the American Society of Health-System Pharmacists (ASHP) released an official position statement,

explaining that due to their "distinct knowledge, skills, and abilities," pharmacists should "lead efforts to guide optimal drug selection and drug dosing based on those results."[111]

On a basic level, clinical pharmacists use patient-specific data to advise physicians on optimal drug therapy for a particular patient. In a 2013 survey of pharmacists working in United States hospitals, 98% of pharmacists provided formal recommendations on dose adjustments, 93% provided drug information, and 92% were responsible for pharmacokinetics monitoring.[112] The interpretation of pharmacogenetic test results into treatment recommendations is just another variation on this theme and fits neatly into pharmacists' preexisting workflows: if a patient's creatinine clearance is below a certain level, decrease the dose of drug X. If their potassium level is too high, switch to drug Y. If they are a CYP2D6 poor metabolizer, avoid drug Z.

The utility of pharmacists in specialized clinical roles has been well-established in areas such as anticoagulation, smoking cessation, transition of care medication review, and disease state management for chronic conditions like asthma, hypertension, hyperlipidemia, and diabetes. [112,113] These programs are effective at improving patient outcomes. As reported in a 2015 meta-analysis on the effectiveness of clinical pharmacy services, anticoagulation programs significantly reduced the risk of total bleeding events (RR 0.51, 95% CI 0.28 – 0.94), hypertension management programs resulted in systolic blood pressure reductions of 8 – 11 mmHg, diabetes management programs resulted in HbA1c reductions of 0.9 – 2.1%, and hyperlipidemia programs resulted in a mean total cholesterol reduction of 22 mg/dL.[113]

3.2.3 Pharmacist roles in existing pharmacogenomics programs

Institutions such as St. Jude Children's Research Hospital and University of Florida have designed pharmacogenomics implementation programs that rely on pharmacist input.

At St. Jude Children's Research Hospital, pharmacists have roles throughout the pharmacogenetic testing process.[114] After diplotype results are available from the laboratory, a pharmacist manually reviews them for each patient and writes a consult note that includes a phenotype assignment, diplotype interpretation, dosing recommendations, activity score (if *CYP2D6* was tested), and a link to online resources for more information. The pharmacist also performs a medication reconciliation to assess whether the patient is taking a medication with a relevant pharmacogenetic association. A second pharmacist reviews all information prior to inclusion in the electronic medical record. All new pharmacist hires at St. Jude receive training in pharmacogenetics as part of their initial required competencies and complete refresher courses every three years. In addition to training on test interpretation and dosing recommendations, pharmacists are also educated on how to discuss pharmacogenetic test results with patients and their families.[114]

University of Florida's Personalized Medicine Program was developed and is run primarily by pharmacists in a wide range of roles.[112] A unique aspect of this program is that whenever a prescriber encounters a point of care alert about a pharmacogenetic association, a clinical pharmacist is notified via an inbox message. This allows the pharmacist to coordinate care in-person with the prescriber or provide direct patient care, if needed. To keep abreast of current pharmacogenetics research, a dedicated pharmacist

is responsible for reviewing new literature, integrating relevant findings into practice, and providing ongoing clinical education. The program also has a PGY2 pharmacy resident who is responsible for monitoring medication safety trends and providing education, data management, and logistical support to prescribers.[112]

3.2.4 Expansion beyond academic medical centers

Currently, most pharmacogenomics implementation programs are based at large academic medical centers. In order to reach patients who don't have access to these facilities, a reasonable next step is implementation in primary care settings. As approximately 80% of primary care visits result in drug prescriptions, pharmacogenetic testing has great potential for widespread benefit.[115] Thus far, very few studies have evaluated the feasibility of using pharmacogenetics in primary care. Preliminary results, however, are promising.

At two family practice offices on Vancouver Island and three family practices and one pharmacy in metropolitan Vancouver, British Columbia, 191 patients were enrolled in a study to assess the ability to obtain DNA samples, genotype them, and link the results to decision support software.[116] Saliva samples were collected using the Oragene DNA collection kit and transported to a laboratory at the University of British Columbia by research staff, mail, or float plane. The DNA was extracted and genotyped using a custom TaqMan panel developed based on genetic variants with the highest levels of clinical evidence for use in primary care. Of the 191 total patients, 185 were successfully genotyped for all genes on the panel. Actionable diplotypes for drugs included in the decision support software were found in 179 (96.8%) of the patients, which is similar to a previous study that found actionable variants in 96% of a 5000 patient cohort. Results

were uploaded into dedicated medication decision support software (MDSS) and reports were distributed directly to each patient's pharmacist or physician. Over a course of 3 months, pharmacists and physicians accessed the software 236 times, most frequently to obtain therapy recommendations for hyperlipidemia (n = 53) and hypertension (n = 52).[116] This study is particularly interesting because it illustrates how the authors addressed the barriers associated with geographically remote clinic sites and electronic healthcare record fragmentation. Instead of excluding patients who lived on Vancouver Island, a system was established to transport samples via float plane. To overcome the informatics challenges of integrating pharmacogenetics data into electronic medical records, software that operates independently of the electronic medical record was developed.

In any setting, implementation of pharmacogenetics requires accurate genotyping data, leadership by knowledgeable healthcare professionals (best embodied by pharmacists), and some creative problem solving. We have all the pieces and now is the time to put them together.

4 Appendix: optimization of human liver microsome incubations

As discussed in section 1.3.2.1, optimization of human liver microsome incubations requires selection of appropriate drug concentration, HLM protein concentration, and incubation time. Since these parameters are interdependent, only one parameter can vary at a time while the others are held constant for optimization assays. For early optimization experiments, a matrix design was used to try several combinations of drug concentrations, protein concentration, and incubation times. Then, specific parameters were explored more in-depth if needed. Pooled human liver microsomes (Xenotech) were used for optimization assays.

4.1 Selection of probe drugs and drug concentrations

Dextromethorphan and metoprolol were selected for use as probe drugs. Dextromethorphan is a classic CYP2D6 probe substrate for *in vitro* assays and has been used as such since the late 1980s.[117,118] CYP2D6 catalyzes the O-demethylation of dextromethorphan to dextrorphan and is responsible for 80% of total dextrorphan formation, with CYP3A4 also contributing when dextromethorphan concentrations are above 50 μM.[119] Metoprolol is not as widely used as a CYP2D6 probe drug but was selected because of the clinical evidence in favor of a pharmacogenetic association. A meta-analysis of 11 clinical trials that included a total of 264 participants found a 5.3-fold difference in peak plasma metoprolol concentration between CYP2D6 poor metabolizers and ultrarapid metabolizers, as well as a 13-fold difference in area under the curve and a

15-fold difference in apparent oral clearance.[120] CYP2D6 is exclusively responsible for the alpha-hydroxylation of metoprolol to alpha-hydroxymetoprolol.[121]

Because a metabolite formation assay approach was selected, probe drug concentrations were required to be below K_m to ensure that the reaction stayed in the linear range of the Michaelis-Menten curve. The concentration cannot be too low, however, because an underlying assumption of the assay is that substrate concentrations must remain high relative to enzyme concentrations. Generally, this assumption holds true if no more than 10% of the substrate is depleted over the course of an incubation.[122]

For each substrate, we searched the literature for the lowest published K_m values, selected based on data compiled in the University of Washington Drug Interaction Database (https://didb.druginteractioninfo.org/). In initial assays, we chose a range of drug concentrations that bracketed these literature K_m values of 2.8 μM for dextromethorphan and 21.6 μM for metoprolol.

4.2 Selection of HLM protein concentrations and incubation times

Typical HLM protein concentrations for incubation assays are in the range of 0.1 – 0.5 mg/mL, but this was not feasible with our limited allocation of HLMs. Based on consultation with our collaborator who quantified the metabolites (section 2.2.2.5), we started with protein concentrations of 50 μg/reaction (0.05 mg/mL) using pooled HLMs and metoprolol at a range of concentrations that bracketed the literature K_m of ~20 μM: 2 μM (1/10 K_m), 4 μM (1/5 K_m), 10 μM (1/2 K_m), 20 μM (K_m), and 40 μM (2 K_m).

In general, incubation times of 5 – 10 minutes will result in linear reactions with detectable metabolites.[123] Because we were limited by low HLM protein concentrations, we tried longer incubation times to ensure the production of sufficient detectable

metabolite. Longer incubations may result in enzyme or substrate degradation or inhibitory effects due to accumulation of metabolites, so we tried a range of incubation times to see if we saw any evidence of these effects at longer timepoints. For the initial metoprolol assay discussed above, incubation times were 5, 15, and 30 minutes.

The metabolite of interest, alpha-hydroxymetoprolol, was detected at every metoprolol concentration and timepoint. At all three timepoints, nonlinearity started to be observed around a metoprolol concentration of 20 μM. Substrate depletion was above 10% for the incubations using 2 μM of metoprolol at 15 and 30 minutes. Thus, acceptable parameters were 2 μM metoprolol + 5 minutes; 4 μM + 5, 10, or 15 minutes; and 10 μM metoprolol + 5, 10, or 15 minutes.

Because we saw ample metabolite formation using 50 μg of pooled HLM protein, we did a similar assay to see if it was feasible to use 20 μg per reaction well. Metoprolol was used at 2 μM and 4 μM at incubation times of 15, 30, and 45 minutes. Longer incubations were tested to account for lower HLM protein concentrations. The rate of alpha-hydroxymetoprolol formation was linear with respect to substrate concentration and time. Substrate depletion remained below 10% for both substrate concentrations at all timepoints. The assay was repeated for confirmation and similar results were obtained.

From these optimization assays, we selected the following parameters to use with the liver bank samples: 20 μg HLM protein, 4 μM metoprolol, and 30 minute incubation time. These parameters produced a linear reaction with 1.14% of substrate depleted over the course of the incubation.

Optimization of incubations using dextromethorphan as a probe drugs were performed similarly. We selected a range of dextromethorphan concentrations that

bracketed the literature K_m (2.8 μM, rounded up to 5 μM to simplify calculations). These concentrations were 0.5 μM (1/10 K_m), 1.25 μM (1/4 K_m), 2.5 μM (1/2 K_m), 5 μM (K_m), and 10 μM (2 K_m). Incubation times were 10, 20, and 30 minutes and each reaction well used 20 μg HLM protein.

The metabolite formation curve appeared linear at metoprolol concentrations below 2.5 μM at every timepoint. The assay was repeated using dextromethorphan at concentrations of 0.5 μM, 1.25 μM, and 2.5 μM and incubation times of 10, 20, and 30 minutes. Dextrorphan formation was linear over time when dextromethorphan was used at concentrations of 0.5 μM and 1.25 μM, but not 2.5 μM.

From these optimization assays, we selected the following parameters to use with the liver bank samples: 20 μg HLM protein, 1.5 μM dextromethorphan, and 20 minute incubation time. Approximately 7% of substrate was depleted over the incubation time with these parameters. Dextromethorphan was used at 1.5 μM rather than 1.25 μM because it was easier to accurately measure the volume of stock solution needed for the assay. A 20 minute incubation time was selected because a longer incubation was not needed to produce quantifiable metabolite.

5 References

1. Pirmohamed M. Pharmacogenetics: past, present and future. *Drug Discovery Today*. 2011;16(19-20):852-861. doi:10.1016/j.drudis.2011.08.006.

2. Zhang G, Nebert DW. Personalized medicine: Genetic risk prediction of drug response. *Pharmacology and Therapeutics*. February 2017:1-16. doi:10.1016/j.pharmthera.2017.02.036.

3. Meyer UA. Pharmacogenetics–five decades of therapeutic lessons from genetic diversity. *Nature Reviews Genetics*. 2004;5(9):669-676. doi:10.1038/nrg1428.

4. Hamburg MA, Collins FS. The path to personalized medicine. *New England Journal of Medicine*. 2010;363(4):301-304. doi:10.1056/NEJMp1006304.

5. Burke W, Thummel K. A call for accurate pharmacogenetic labeling. *JAMA Intern Med*. 2014. doi:10.1001/jamainternmed.2014.5266.

6. Wang B, Canestaro WJ, Choudhry NK. Clinical Evidence Supporting Pharmacogenomic Biomarker Testing Provided in US Food and Drug Administration Drug Labels. *JAMA Intern Med*. 2014;174(12):1938-1939. doi:10.1001/jamainternmed.2014.5266.

7. Sauver J, Bielinski SJ, Olson JE, Bell EJ. Integrating pharmacogenomics into clinical practice: promise vs reality. *The American Journal of …*. 2016. doi:10.1016/j.amjmed.2016.04.009.

8. O'Donnell PH, Danahey K, Jacobs M, et al. Adoption of a clinical pharmacogenomics implementation program during outpatient care-initial results of the University of Chicago "1,200 Patients Project." Williams MS, ed. *Am J Med Genet*. 2014;166(1):68-75. doi:10.1002/ajmg.c.31385.

9. CPIC: Clinical Pharmacogenetics Implementation Consortium. pharmgkb.org. https://www.pharmgkb.org/page/cpic. Accessed May 13, 2016.

10. Johnson JA, Caudle KE, Gong L, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. *Clin Pharm Ther*. 2017;102(3):397-404. doi:10.1002/cpt.668.

11. Caudle KE, Klein TE, drug JHC, 2014. Incorporation of pharmacogenomics into routine clinical practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline development process. *ingentaconnectcom*.15(2):209-217.

12. Yang L, Price ET, Chang C-W, et al. Gene Expression Variability in Human

Hepatic Drug Metabolizing Enzymes and Transporters. Miao X-P, ed. *PLoS ONE*. 2013;8(4):e60368–12. doi:10.1371/journal.pone.0060368.

13. Gates R. Phase I Drug Metabolism. Sigma-Aldrich. http://www.sigmaaldrich.com/technical-documents/articles/biofiles/phase-i-drug-metabolism.html. Accessed May 18, 2016.

14. Zanger UM, Schwab M. Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacology and Therapeutics*. 2013;138(1):103-141. doi:10.1016/j.pharmthera.2012.12.007.

15. Jancova P, Anzenbacher P, Anzenbacherova E. Phase II drug metabolizing enzymes. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2010;154(2):103-116.

16. Leschziner GD, Andrew T, Pirmohamed M, Johnson MR. ABCB1 genotype and PGP expression, function and therapeutic drug response: a critical review and recommendations for future research. *Pharmacogenomics J*. 2006;7(3):154-179. doi:10.1038/sj.tpj.6500413.

17. Shugarts S, Benet LZ. The Role of Transporters in the Pharmacokinetics of Orally Administered Drugs. *Pharm Res*. 2009;26(9):2039-2054. doi:10.1007/s11095-009-9924-0.

18. Bruhn O, Cascorbi I. Polymorphisms of the drug transporters ABCB1, ABCG2, ABCC2 and ABCC3 and their impact on drug bioavailability and clinical relevance. *Expert Opinion on Drug Metabolism & Toxicology*. 2014;10(10):1337-1354. doi:10.1517/17425255.2014.952630.

19. Ramsey LB, Johnson SG, Caudle KE, et al. The Clinical Pharmacogenetics Implementation Consortium Guideline for SLCO1B1 and Simvastatin-Induced Myopathy: 2014 Update. *Clin Pharmacol Ther*. 2014;96(4):423-428. doi:10.1038/clpt.2014.125.

20. Lee HH, Ho RH. Interindividual and interethnic variability in drug disposition: polymorphisms in organic anion transporting polypeptide 1B1 (OATP1B1; SLCO1B1). *Br J Clin Pharmacol*. 2017;83(6):1176-1184. doi:10.1111/bcp.13207.

21. Relling MV, Evans WE. Pharmacogenomics in the clinic. *Nature*. 2015;526(7573):343-350. doi:10.1038/nature15817.

22. Yang W, Wu G, Broeckel U, et al. Comparison of genome sequencing and clinical genotyping for pharmacogenes. *Clin Pharmacol Ther*. June 2016:1-30. doi:10.1002/cpt.411.

23. Drug Labels. pharmgkb.org. https://www.pharmgkb.org/view/drug-labels.do. Accessed June 30, 2016.

24. Drug Label Information and Legend. pharmgkb.org. https://www.pharmgkb.org/page/drugLabelLegend. Accessed June 30, 2016.

25. Tarceva drug label. NIH DailyMed. https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=57bccb29-1c47-4c64-ab6a-77960a91cc20. Published June 14, 2016. Accessed June 30, 2016.

26. Herceptin drug label. NIH DailyMed. https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=492dbdb2-077e-4064-bff3-372d6af0a7a2. Published April 30, 2016. Accessed June 30, 2016.

27. Abacavir drug label. https://www.pharmgkb.org/view/drug-label.do?id=PA166104833. Published March 1, 2016. Accessed June 30, 2016.

28. Carbamazepine drug label. NIH DailyMed. https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=c13bc0b8-7900-4ef4-98ed-e1315a08d95d. Published October 18, 2013. Accessed June 30, 2016.

29. PLAVIX (clopidogrel bisulfate) tablets. NIH DailyMed. https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=2ed86bc8-8ea5-4ffa-a762-4fc266c1e620. Published February 2, 2017. Accessed June 30, 2016.

30. Hoffman JM, Dunnenberger HM, Kevin Hicks J, et al. Developing knowledge resources to support precision medicine: principles from the Clinical Pharmacogenetics Implementation Consortium (CPIC). *J Am Med Inform Assoc*. March 2016:ocw027–6. doi:10.1093/jamia/ocw027.

31. CPIC Guidelines. CPIC. https://cpicpgx.org/guidelines/. Accessed May 15, 2017.

32. Dunnenberger HM, Crews KR, Hoffman JM, et al. Preemptive Clinical Pharmacogenetics Implementation: Current Programs in Five US Medical Centers. *Annu Rev Pharmacol Toxicol*. 2015;55(1):89-106. doi:10.1146/annurev-pharmtox-010814-124835.

33. Crews KR. Clinical Implementation of Pharmacogenomics Through a Pharmacist-Managed Service. In:; 2012:1-36.

34. Scott SA, Sangkuhl K, Stein CM, et al. Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C19 Genotype and Clopidogrel Therapy: 2013 Update. *Clin Pharmacol Ther*. 2013;94(3):317-323. doi:10.1038/clpt.2013.105.

35. Guengerich FP. Cytochrome P450 and Chemical Toxicology. *Chem Res Toxicol*. 2008;21(1):70-83. doi:10.1021/tx700079z.

36. Krishnan S, Schenkman JB, Rusling JF. Bioelectronic Delivery of Electrons to Cytochrome P450 Enzymes. *J Phys Chem B*. 2011;115(26):8371-8380. doi:10.1021/jp201235m.

37. de Montellano PRO. *Cytochrome P450: Structure, Mechanism, and Biochemistry*. Springer; 2015.

38. Munro AW, Girvan HM, Mason AE, Dunford AJ, McLean KJ. What makes a P450 tick? *Trends in Biochemical Sciences*. 2013;38(3):140-150. doi:10.1016/j.tibs.2012.11.006.

39. Poulos TL, Johnson EF. Structure of Cytochrome P450 Enzymes. In: Ortiz de Montellano PR, ed. *Cytochrome P450: Structure, Mechanism, and Biochemistry*. Springer International Publishing; 2015:3-32.

40. Reynald RL, Sansen S, Stout CD, Johnson EF. Structural Characterization of Human Cytochrome P450 2C19. *Journal of Biological Chemistry*. 2012;287(53):44581-44591. doi:10.1074/jbc.M112.424895.

41. Obach RS. Measurement of Michaelis Constants for Cytochrome P450-Mediated Biotransformation Reactions Using a Substrate Depletion Approach. *Drug Metabolism and Disposition*. 2002;30(7):831-837. doi:10.1124/dmd.30.7.831.

42. Wienkers LC, Stevens JC. Cytochrome P450 Reaction Phenotyping. In: Lee JS, Obach RS, Fisher MB, eds. *Drug Metabolizing Enzymes*. 2003:255-310.

43. Houston JB, Kenworthy KE, Galetin A. Typical and Atypical Enzyme Kinetics. In: *Drug Metabolizing Enzymes*. Marcel Dekker, Inc; 2003:211-254.

44. Fasinu P, Bouic PJ, Rosenkranz B. Liver-based in vitro technologies for drug biotransformation studies-a review. *Current Drug Metabolism*. 2012;13(2):215-224. doi:10.2174/138920012798918426.

45. Asha S, Vidyavathi M. Role of Human Liver Microsomes in In Vitro Metabolism of Drugs—A Review. *Appl Biochem Biotechnol*. 2009;160(6):1699-1722. doi:10.1007/s12010-009-8689-6.

46. Human Liver Microsomes. xenotech.com. https://www.xenotech.com/products/subcellular-fractions/human/liver/microsomes. Accessed July 10, 2016.

47. Jones HM, Houston JB. Substrate depletion approach for determining in vitro metabolic clearance: time dependencies in hepatocyte and microsomal incubations. *Drug Metabolism and Disposition*. 2004. doi:10.1124/dmd.104.000125.

48. Definition: allele. nature.com. http://www.nature.com/scitable/definition/allele-48. Accessed June 26, 2016.

49. Definition: haplotype/haplotypes. nature.com. http://www.nature.com/scitable/definition/haplotype-142. Accessed June 26, 2016.

50. Sim SC, Ingelman-Sundberg M. The Human Cytochrome P450 (CYP) Allele Nomenclature website: a peer-reviewed database of CYP variants and their associated effects. *Human Genomics 2010 4:4*. 2010;4(4):1. doi:10.1186/1479-7364-4-4-278.

51. Woolf SH. The Meaning of Translational Research and Why It Matters. *JAMA*. 2008;299(2):211-213. doi:10.1001/jama.2007.26.

52. Gaedigk A, Simon SD, Pearce RE. The CYP2D6 activity score: translating genotype information into a qualitative measure of phenotype. *Clin Pharmacol Ther*. 2008;83(2):234-242. doi:10.1038/sj.clpt.6100406.

53. Rae JM, Drury S, Hayes DF, et al. CYP2D6 and UGT2B7 Genotype and Risk of Recurrence in Tamoxifen-Treated Breast Cancer Patients. *JNCI Journal of the National Cancer Institute*. 2012;104(6):452-460. doi:10.1093/jnci/djs126.

54. Regan MM, Leyland-Jones B, Bouzyk M, et al. CYP2D6 Genotype and Tamoxifen Response in Postmenopausal Women with Endocrine-Responsive Breast Cancer: The Breast International Group 1-98 Trial. *JNCI Journal of the National Cancer Institute*. 2012;104(6):441-451. doi:10.1093/jnci/djs125.

55. Nakamura Y, Ratain MJ, Cox NJ. Re: CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the Breast International Group 1-98 trial. *Journal of the National Cancer Institute*. 2012;104(16):1264. doi:10.1093/jnci/djs312.

56. Pratt VM, McLeod HL, Dean L, Malheiro A, Rubinstein W. *Medical Genetics Summaries*. 2017:1-405.

57. Koren G, Cairns J, Chitayat D, Gaedigk A, Leeder SJ. Pharmacogenetics of morphine poisoning in a breastfed neonate of a codeine-prescribed mother. *The Lancet*. 2006;368(9536):704–1. doi:10.1016/S0140-6736(06)69255-6.

58. Postmarket Drug Safety Information for Patients and Providers - Use of Codeine Products in Nursing Mothers - Questions and Answers. FDA CDER. http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm118113.htm. Accessed June 7, 2016.

59. Crews KR, Gaedigk A, Dunnenberger HM, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for Codeine Therapy in the Context of Cytochrome P450 2D6 (CYP2D6) Genotype. *Clin Pharmacol Ther*. 2009;91(2):321-326. doi:10.1038/clpt.2011.287.

60. Hicks JK, Swen JJ, Thorn CF, et al. Clinical Pharmacogenetics Implementation Consortium Guideline for CYP2D6 and CYP2C19 Genotypes and Dosing of Tricyclic Antidepressants. *Clin Pharmacol Ther*. 2013;93(5):402-408. doi:10.1038/clpt.2013.2.

61. Hicks JK, Bishop JR, Sangkuhl K, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6and CYP2C19Genotypes and Dosing of Selective Serotonin Reuptake Inhibitors. *Clin Pharmacol Ther*. 2015;98(2):127-134. doi:10.1002/cpt.147.

62. Genes-Drugs. CPIC. https://cpicpgx.org/genes-drugs/. Accessed May 15, 2017.

63. NUEDEXTA® (dextromethorphan hydrobromide and quinidine sulfate) capsules. PharmGKB. https://www.pharmgkb.org/view/drug-label.do?id=PA166104886. Published March 27, 2012. Accessed May 15, 2017.

64. Gaedigk A. Complexities of CYP2D6 gene analysis and interpretation. *International Review of Psychiatry*. 2013;25(5):534-553. doi:10.3109/09540261.2013.825581.

65. Teh LK, Bertilsson L. Pharmacogenomics of CYP2D6: Molecular Genetics, Interethnic Differences and Clinical Importance. *Drug Metabolism and Pharmacokinetics*. 2012;27(1):55-67. doi:10.2133/dmpk.DMPK-11-RV-121.

66. Yasukochi Y, Satta Y. Molecular Evolution of the CYP2D Subfamily in Primates: Purifying Selection on Substrate Recognition Sites without the Frequent or Long-Tract Gene Conversion. *Genome Biology and Evolution*. 2015;7(4):1053-1067. doi:10.1093/gbe/evv056.

67. Gaedigk A, Bhathena A, Ndjountché L, et al. Homo sapiens cytochrome P4502D6 (CYP2D6) gene, CYP2D6*1 allele, complete cds. *Pharmacogenomics J*. 2005;5(3):173-182. doi:10.1038/sj.tpj.6500305.

68. Kimura S, Umeno M, Skoda RC, Meyer UA, Gonzalez FJ. The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene. *American Journal of Human Genetics*. 1989;45(6):889.

69. Hanioka N, Kimura S, Meyer UA. The human CYP2D locus associated with a common genetic defect in drug oxidation: a G1934----A base change in intron 3 of a mutant CYP2D6 allele results in an aberrant 3' splice recognition site. *American Journal of Human Genetics*. 1990.

70. *CYP2D6* allele nomenclature. The Human Cytochrome P450 (CYP) Allele Nomenclature Database. http://www.cypalleles.ki.se/cyp2d6.htm. Accessed April 12, 2016.

71. Zhou S. *Cytochrome P450 2D6: Structure, Function, Regulation, and Polymorphism*. 2017:1-532.

72. CYP2D6*2 VIP Annotation. pharmgkb.org. https://www.pharmgkb.org/haplotype/PA165816577#tabview=tab3&subtab=. Accessed June 13, 2016.

73. Baldi P, Long AD. A Bayesian framework for the analysis of microarray expression data: Regularized t-test and statistical inferences of gene changes. *Bioinformatics*. 2001;17(6):509-519. doi:10.1093/bioinformatics/17.6.509.

74. Ingelman-Sundberg M, Sim SC. Intronic polymorphisms of cytochromes P450. *Human Genomics 2010 4:4*. 2010;4(6):402-405.

75. Black JL, Walker DL, O'Kane DJ, Harmandayan M. Frequency of Undetected CYP2D6 Hybrid Genes in Clinical Samples: Impact on Phenotype Prediction. *Drug Metabolism and Disposition*. 2011;40(1):111-119. doi:10.1124/dmd.111.040832.

76. Gaedigk A, Jaime LKM, Bertino JS, et al. Identification of Novel CYP2D7-2D6 Hybrids: Non-Functional and Functional Variants. *Front Pharmacol*. 2010;1:1-9. doi:10.3389/fphar.2010.00121.

77. Saxena R, Shaw GL, Relling MV, et al. Identification of a new variant CYP2D6 allele with a single base deletion in exon 3 and its association with the poor metabolizer phenotype. *Human Molecular Genetics*. 1994;3(6):923-926. doi:10.1093/hmg/3.6.923.

78. Tyndale R, Aoyama T, Broly F, et al. Identification of a new variant CYP2D6 allele lacking the codon encoding Lys-281: possible association with the poor metabolizer phenotype. *Pharmacogenetics and Genomics*. 1991;1(1):26.

79. Cai J, Dai D-P, Geng P-W, et al. Effects of 22 Novel CYP2D6 Variants Found in the Chinese Population on the Bufuralol and Dextromethorphan Metabolisms In Vitro. *Basic & Clinical Pharmacology & Toxicology*. 2015;118(3):190-199. doi:10.1111/bcpt.12478.

80. Oscarson M, Hidestrand M, Johansson I, Ingelman-Sundberg M. A Combination of Mutations in the CYP2D6*17(CYP2D6Z) Allele Causes Alterations in Enzyme Function. *Mol Pharmacol*. 1997;52(6):1034-1040. doi:10.1124/mol.52.6.1034.

81. Marez-AIlorge D, EHis SW, Guidice J-ML, Tucker GT, Broly E. A rare G2061 insertion affecting the open reading frame of CYP2D6 and responsible for the poor metabolizer phenotype. *Pharmacogenetics and Genomics*. 1999;9(3):393-396.

82. Wennerholm A, Johansson I, Hidestrand M, Bertilsson L, Gustafsson LL, Ingelman-Sundberg M. Characterization of the CYP2D6*29 allele commonly present in a black Tanzanian population causing reduced catalytic activity. *Pharmacogenetics and Genomics*. 2001;11(5):417.

83. RAIMUNDO S. A novel intronic mutation, 2988G>A, with high predictivity for impaired function of cytochrome P450 2D6 in white subjects*1. *Clin Pharmacol Ther*. 2004;76(2):128-138. doi:10.1016/j.clpt.2004.04.009.

84. Toscano C, Klein K, Blievernicht J, et al. Impaired expression of CYP2D6 in intermediate metabolizers carrying the *41 allele caused by the intronic SNP 2988G>A: evidence for modulation of splicing events. *Pharmacogenetics and Genomics*. 2006;16(10):755-766. doi:10.1097/01.fpc.0000230112.96086.e0.

85. Gaedigk A, Sangkuhl K, Whirl Carrillo M, Klein T, Leeder JS. Prediction of CYP2D6 phenotype from genotype across world populations. *Genet Med*. 2016;19(1):69-76. doi:10.1038/gim.2016.80.

86. Black JL, Walker DL, O'Kane DJ, Harmandayan M. Frequency of Undetected CYP2D6 Hybrid Genes in Clinical Samples: Impact on Phenotype Prediction. *Drug Metabolism and Disposition*. 2011;40(1):111-119. doi:10.1124/dmd.111.040832.

87. Sistonen J, Sajantila A, Lao O, Corander J, Barbujani G, Fuselli S. CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. *Pharmacogenetics and Genomics*. 2007;17(2):93. doi:10.1097/01.fpc.0000239974.69464.f2.

88. Bebia Z, Buch SC, Wilson JW, Frye R, 2004. Bioequivalence revisited: influence of age and sex on CYP enzymes. *Clinical Pharmacology and Therapeutics*. doi:10.1016/j.clpt.2004.08.021.

89. Center for Drug Evaluation, Research. Drug Interactions & Labeling - Drug Development and Drug Interactions: Table of Substrates, Inhibitors and Inducers.

90. Sandee D, Morrissey K, Agrawal V, et al. Effects of genetic variants of human P450 oxidoreductase on catalysis by CYP2D6 in vitro. *Pharmacogenetics and Genomics*. 2010;20(11):677-686. doi:10.1097/FPC.0b013e32833f4f9b.

91. Chaudhry AS, Thirumaran RK, Yasuda K, et al. Genetic Variation in Aldo-Keto Reductase 1D1 (AKR1D1) Affects the Expression and Activity of Multiple Cytochrome P450s. *Drug Metabolism and Disposition*. 2013;41(8):1538-1547. doi:10.1124/dmd.113.051672.

92. Yang Y, Botton MR, Scott ER, Scott SA. Sequencing the CYP2D6 gene: from variant allele discovery to clinical pharmacogenetic testing. *Pharmacogenomics*. 2017. doi:10.2217/pgs-2017-0033.

93. Gordon AS, Fulton RS, Qin X, Mardis ER. PGRNseq: a targeted capture sequencing panel for pharmacogenetic research and implementation. 2016;26(4):161-168. doi:10.1097/FPC.0000000000000202.

94. Gardiner SJ. Pharmacogenetics, Drug-Metabolizing Enzymes, and Clinical Practice. *Pharmacological Reviews*. 2006;58(3):521-590. doi:10.1124/pr.58.3.6.

95. Gaedigk A, Riffel AK, Berrocal BG, Solaesa VG, Dávila I, Isidoro-García M. Characterization of a complex CYP2D6 genotype that caused an AmpliChip

CYP450 Test® no-call in the clinical setting. *Clinical Chemistry and Laboratory Medicine (CCLM)*. 2014;52(6):1-9. doi:10.1515/cclm-2013-0943.

96. Table of Pharmacogenomic Biomarkers in Drug Labeling. NIH CDER. https://www.fda.gov/downloads/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/UCM545881.pdf. Published March 13, 2017. Accessed May 12, 2017.

97. CYP2D6. PharmGKB. https://www.pharmgkb.org/gene/PA128. Accessed May 15, 2017.

98. Shirasaka Y, Chaudhry AS, McDonald M, et al. Interindividual variability of CYP2C19-catalyzed drug metabolism due to differences in gene diplotypes and cytochrome P450 oxidoreductase content. *Pharmacogenomics J*. September 2015:1-13. doi:10.1038/tpj.2015.58.

99. Prasad B, Unadkat JD. Optimized Approaches for Quantification of Drug Transporters in Tissues and Cells by MRM Proteomics. *AAPS J*. 2014;16(4):634-648. doi:10.1208/s12248-014-9602-y.

100. Tanner JA, Prasad B, Claw KG, et al. Predictors of Variation in CYP2A6 mRNA, Protein, and Enzyme Activity in a Human Liver Bank: Influence of Genetic and Nongenetic Factors. *Journal of Pharmacology and Experimental Therapeutics*. 2016;360(1):129-139. doi:10.1124/jpet.116.237594.

101. Crews KR, Gaedigk A, Dunnenberger HM, et al. Clinical Pharmacogenetics Implementation Consortium Guidelines for Cytochrome P450 2D6 Genotype and Codeine Therapy: 2014 Update. *Clin Pharmacol Ther*. 2014;95(4):376-382. doi:10.1038/clpt.2013.254.

102. Kiyotani K, Shimizu M, Kumai T, Kamataki T, Kobayashi S, Yamazaki H. Limited effects of frequent CYP2D6*36-*10 tandem duplication allele on in vivo dextromethorphan metabolism in a Japanese population. *Eur J Clin Pharmacol*. 2010;66(10):1065-1068. doi:10.1007/s00228-010-0876-4.

103. Gaedigk A, Bradford LD, Alander SW, Leeder JS. CYP2D6*36 GENE ARRANGEMENTS WITHIN THE CYP2D6 LOCUS: ASSOCIATION OF CYP2D6*36 WITH POOR METABOLIZER STATUS. 2006;34(4):563-569. doi:10.1124/dmd.105.008292.

104. Koussounadis A, Langdon SP, Um IH, Harrison DJ, Smith VA. Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. *Sci Rep*. June 2015:1-9. doi:10.1038/srep10775.

105. Beoris M, Amos Wilson J, Garces JA, Lukowiak AA. CYP2D6 copy number distribution in the US population. *Pharmacogenetics and Genomics*. 2016;26(2):96-99. doi:10.1097/FPC.0000000000000188.

106. Klein K. Pharmacogenomics of cytochrome P450 3A4: recent progress toward the

"missing heritability" problem. *Frontiers in Genetics*. 2017;4(12):1-15. doi:10.3389/fgene.2013.00012/abstract.

107. Johansen Taber K, Dickinson B. Pharmacogenomic knowledge gaps and educational resource needs among physicians in selected specialties. *PGPM*. July 2014:145–18. doi:10.2147/PGPM.S63715.

108. Haga SB, Burke W, Ginsburg GS, Mills R, Agans R. Primary care physicians' knowledge of and experience with pharmacogenetic testing. *Clin Genet*. 2012;82(4):388-394. doi:10.1111/j.1399-0004.2012.01908.x.

109. Unertl KM, Field JR, Price L, Peterson JF. Clinician perspectives on using pharmacogenomics in clinical practice. *Personalized Medicine*. 2015;12(4):339-347. doi:10.2217/pme.15.10.

110. Weitzel KW, Aquilante CL, Johnson S, Kisor DF, Empey PE. Educational strategies to enable expansion of pharmacogenomics-based care. *American Journal of Health-System Pharmacy*. 2016;73(23):1986-1998. doi:10.2146/ajhp160104.

111. ASHP. ASHP Statement on the Pharmacist's Role in Clinical Pharmacogenomics. *American Journal of Health-System Pharmacy*. 2017;72:579-581.

112. Owusu-Obeng A, Weitzel KW, Hatton RC, et al. Emerging Roles for Pharmacists in Clinical Implementation of Pharmacogenomics. *Pharmacotherapy*. 2014;34(10):1102-1112. doi:10.1002/phar.1481.

113. Rotta I, Salgado TM, Silva ML, Correr CJ, Fernandez-Llimos F. Effectiveness of clinical pharmacy services: an overview of systematic reviews (2000–2010). *International Journal of Clinical Pharmacy*. 2015;37(5):687-697. doi:10.1007/s11096-015-0137-9.

114. Hoffman JM, Haidar CE, Wilkinson MR, et al. PG4KDS: A model for the clinical implementation of pre-emptive pharmacogenetics. Williams MS, ed. *Am J Med Genet*. 2014;166(1):45-55. doi:10.1002/ajmg.c.31391.

115. Haga SB, Allen LaPointe NM, Moaddeb J, Mills R, Patel M, Kraus WE. Pilot study: incorporation of pharmacogenetic testing in medication therapy management services. *Pharmacogenomics*. 2014;15(14):1729-1737. doi:10.2217/pgs.14.118.

116. Dawes M, Aloise MN, Ang JS, et al. Introducing pharmacogenetic testing with clinical decision support into primary care: a feasibility study. *CMAJ Open*. 2016;4(3):E528-E534. doi:10.9778/cmajo.20150070.

117. Dayer P, Leemann T, Striberni R. Dextromethorphan O-demethylation in liver microsomes as a prototype reaction to monitor cytochrome P-450 db1 activity. *Clinical Pharmacology and Therapeutics*. 1989;45(1):34-40.

doi:10.1038/clpt.1989.6.

118. Kronbach T. Bufuralol, Dextromethorphan, and Debrisoquine as Prototype Substrates for Human P450IID6. *Methods in Enzymology*. 1991;206:509-517.

119. Yu A, Haining RL. Comparative Contribution to Dextromethorphan Metabolism by Cytochrome P450 Isoforms in Vitro: Can Dextromethorphan Be Used as a Dual Probe for Both CYP2D6 and CYP3A Activities? *Drug Metabolism and Disposition*. 2001;29(11):1514-1520.

120. Blake CM, Kharasch ED, Schwab M, Nagele P. A Meta-Analysis of CYP2D6 Metabolizer Phenotype and Metoprolol Pharmacokinetics. *Clin Pharmacol Ther*. 2013;94(3):394-399. doi:10.1038/clpt.2013.96.

121. Hamelin BA, Bouayad A, Methot J, et al. Significant interaction between the nonprescription antihistamine diphenhydramine and the CYP2D6 substrate metoprolol in healthy men with high or low CYP2D6 …. *Clinical Pharmacology and Therapeutics*. 2000;67(5):466-477. doi:10.1067/mcp.2000.106464.

122. Brooks HB, Geeganage S, Kahl SD, Montrose C. Basics of enzymatic assays for HTS. In: *Assay Guidance Manual*. 2012.

123. Lee J, Obach RS, Fisher MB. *Drug Metabolizing Enzymes*. Marcel Dekker; 2003.