Planning and Conducting a Pharmacogenetics Association Study

[1]Daniel L Hertz, PharmD, PhD, DLHertz@med.umich.edu

[2]Meghan J. Arwood, PharmD, MS, BCPS, marwood@trhc.com

[3]Gabriele Stocco, stoccog@units.it

[4]Sonal Singh, sonal.singh@takeda.com

[5,6]Jason H. Karnes, PharmD, PhD, BCPS, FAHA karnes@pharmacy.arizona.edu

[7]Laura Ramsey, PhD, Laura.Ramsey@cchmc.org

Affiliations:

[1]Department of Clinical Pharmacy, University of Michigan College of Pharmacy, Ann Arbor, MI, United States, 48109-1065

[2]Tabula Rasa HealthCare, Precision Pharmacotherapy Research and Development Institute, Orlando, FL

[3]Department of Life Sciences, University of Trieste, Italy

[4]Takeda, San Diego, CA, United States, 92121

[5]Department of Pharmacy Practice and Science, University of Arizona College of Pharmacy, Tucson, AZ

[6] Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN

[7]Divisions of Clinical Pharmacology & Research in Patient Services, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH United States, 45229

Corresponding Author:

Daniel L Hertz

428 Church St.

Room 3054 College of Pharmacy

Ann Arbor, MI 48109-1065

Office phone: (734) 763-0015

Fax: (734) 763-4480

E-mail: DLHertz@med.umich.edu

Conflict of interest statement: The authors declared no competing interests for this work.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/CPT.2270

Funding: No funding was received for this work.

Abstract

Pharmacogenetics (PGx) association studies are used to discover, replicate, and validate the association between an inherited genotype and a treatment outcome. The objective of this tutorial is to provide trainees and novice PGx researchers with an overview of the major decisions that need to be made when designing and conducting a PGx association study. The first critical decision is to determine whether the objective of the study is discovery, replication, or validation. Next, the researcher must identify a patient cohort that has all of the data necessary to conduct the intended analysis. Then, the investigator must select and define the treatment outcome, or phenotype, that will be analyzed. Next, the investigator must determine what genotyping approach and genetic data will be included in the analysis. Finally, the association between the genotype and phenotype is tested using some statistical analysis methodology. This tutorial is divided into 5 sections, each section describes commonly used approaches and provides suggestions and resources for designing and conducting a PGx association study. Successful PGx association studies are necessary to discover and validate associations between inherited genetic variation and treatment outcomes, which enable clinical translation to improve efficacy and reduce toxicity of treatment.

Introduction to Pharmacogenetics

The field of pharmacogenetics (PGx) investigates the influence of inherited variants in the patient's germline genome with pharmacotherapeutic outcomes. PGx studies are conducted along the translational research spectrum, from initial discovery of an association between a genetic variant and an outcome to implementation studies determining how best to integrate PGx testing into clinical care. The initial steps of discovering and validating the association between a genotype and outcome, which we will refer to as PGx association studies. These studies are commonly conducted by individuals without formal training in PGx methods, who would benefit from basic guidelines describing the general principles of PGx association studies.

The objective of this tutorial is to introduce trainees and novice investigators to the general process of PGx association studies. This process is divided into the five main considerations when designing a PGx

association study; study objective, patient cohort, phenotype, genotype, and statistical testing (**Figure 1**). We describe and provide helpful recommendations for each of these five major steps. Although we introduce some basic concepts in clinical study design that are not specific to PGx, individuals who are unfamiliar with these fundamental topics should first consult an introductory review.[1] This tutorial is intended for individuals with an understanding of basic clinical research who are relatively new to PGx; it is not intended to provide a comprehensive review of all strategies for PGx association studies and topics that are most relevant to advanced PGx researchers may not be discussed or mentioned only briefly.

1. Study Objective

The first determination that needs to be made for any research study, including a PGx analysis, is the study objective. As mentioned earlier, PGx studies span the translational research spectrum from discovery through implementation. PGx research begins with the discovery of a putative association between an inherited genetic variant (genotype) and a clinical outcome (phenotype), which should then be replicated and validated in independent patient cohorts. Confirmation of a PGx association is referred to as "clinical validity,"[2] which is necessary but not sufficient for translating PGx into clinical practice. Clinical translation usually requires demonstration of "clinical utility," meaning that genetics-informed treatment improves clinical outcomes. Clinical utility is typically demonstrated in prospective clinical trials comparing genotype-directed care with standard of care treatment. Clinical translation and implementation will not be discussed in this tutorial but have been reviewed elsewhere.[3, 4]

This tutorial will focus on PGx analyses to discover, replicate, or validate associations. It is necessary to determine which of these is your study objective to guide selection of an appropriate cohort, phenotype, genotype, and statistical analysis, as discussed within each section of this tutorial. The objective of a PGx discovery study is to identify a novel PGx association for future replication and validation. For that reason, discovery studies seek to maximize the likelihood of detecting associations. To achieve this goal, discovery studies often test many potential genetic associations with liberal statistical procedures, leading to many discovery-phase associations that are possibly false positives.[5-7] It is rarely, if ever, appropriate to take a discovery-phase association and attempt to translate it into clinical practice. Instead, discovery-phase PGx associations must be successfully replicated with similar direction of effect in several independent patient cohorts. These replication studies also often test several previously discovered associations without strict statistical analysis procedures. The objective of replication studies is to determine whether the PGx discoveries can be replicated (i.e., are they likely to

be true associations) and how robust the association is when tested in slightly different cohorts with slightly different phenotypes. The final phase is validation of the association between the genotype and phenotype. Validation can be accomplished through consistent, successful replication in multiple independent cohorts or via a single well-conducted validation study using a single, pre-specified genetic predictor, phenotype definition, and statistical analysis plan. Upon clinical validation, a PGx association is likely ready for prospective testing to demonstrate clinical utility and justify clinical implementation.

2. Patient Cohort

Ethics and Regulatory Oversight

Conducting PGx association studies requires access to genetic and clinical data, which may involve collection and analysis of DNA and personal health information.[8] Data and sample collection and analyses must be performed following the general ethical principles for human subjects research, as described in the Declaration of Helsinki.[9] These principles require that human subjects are aware of and consent to experiments in which they are a participant and require experimenters to take necessary precautions to protect participants' safety and confidentiality. Although PGx association studies have limited direct risks for participants, the permanence of genetic information and possibility that it can be used to predict medical outcomes leaves it vulnerable to being used for discriminatory purposes. In the United States, regulations such as the Genetic Information Nondiscrimination Act have limited the potential negative impact of collecting patients' genetic data.[10] However, care should always be taken to protect patient samples and information, such as anonymization or de-identification.[11]

Prior to collecting any samples or data for a PGx analysis, it is critical that the study be reviewed by an Institutional Review Board (IRB) and/or Ethics Board. These committees evaluate the soundness, relevance, and appropriateness of the scientific question, study design, and procedures to obtain, collect, store, and analyze patient information. The necessary approval will depend on the institution and study design. PGx association studies are often classified as having no more than minimal risk to patients and may be approved by expedited review. Alternatively, IRB approval may not be necessary if it is determined that the study satisfies criteria for an exemption to human subject's research. This is often the case when using linked genetic and health information that was previously collected within an IRB-approved research study,[12] which satisfies the criteria for secondary use of existing data and samples.[13]

Retrospective and Prospective Cohorts

PGx association studies can be conducted using patient data that were initially collected retrospectively or prospectively. In a retrospectively collected cohort, the patients have already received treatment and the outcome of interest has already occurred at the time the study is initiated.[14]. Data are typically abstracted from the patients' medical record or an existing database, which is much less time consuming and costly than prospective data collection. However, this real-world data is likely collected for a purpose other than clinical research, and there may be a lot of important data that is missing, sporadically collected, or inaccurate.[15] Collecting unstructured data from abstracting clinical notes from the electronic medical record is particularly challenging and time consuming, but is often the only feasible strategy to collect the necessary data. Thus, it is typically necessary to perform data pre-processing, in addition to manually cleaning the data, to ensure that only relevant and informative patients are included in the analysis and outcomes are accurately characterized, as described in later sections of this tutorial.

In prospective cohorts the outcome has not occurred at the time the study is initiated and data are collected in real time.[14] Advantages of this strategy include the ability to dictate which patients, treatments, outcomes, and other clinical data are collected. However, prospective data collection is much more time and resource intensive, particularly for infrequent conditions or outcomes. Prospective cohorts can be collected within interventional clinical trials or observational studies, which are differentiated based on whether the protocol dictates the patient's treatment or not, respectively.[16] Clinical trials often enroll relatively large cohorts of homogeneous patients receiving strictly defined treatments from whom outcomes are systematically collected, making these ideal for PGx association testing.[17] Observational studies include registries that are linked to available genetic samples or data collected at the institutional[12] or national[18] level, providing large patient cohorts that are relatively heterogenous in terms of disease and treatment. Large registries are also well suited for PGx discovery or testing whether previously discovered associations are sufficiently robust to be replicated in cohorts of patients that are not as strictly defined and characterized.

Inclusion and Exclusion Criteria

The next important consideration is to determine which potential patients should be included in your analysis.[19] To be considered for inclusion in a PGx association study, a patient usually must have been exposed to the drug of interest, assessed for the outcome of interest, and have provided a biospecimen for genetic analysis. It is sometimes advisable to exclude patients who were treated with an insufficient

number of doses, either due to discontinuation or lack of adherence,[20] but determination of an appropriate threshold will require clinical judgement. The second critical inclusion criterion is assessment of the outcome of interest, or phenotype. Phenotype collection and definition are discussed at length in the Phenotype section of this tutorial. An important point is that patients should be excluded if they were lost to follow up before the outcome of interest could have occurred or if, for any other reason, the available data does not allow you to determine whether or not the patient experienced the outcome.[21] If the outcome is unclear for an individual patient, it is typically better to exclude them from the analysis than to include risk misclassifying them in the analysis. The final necessary inclusion criterion is the availability of a DNA sample for genetic analysis or existing genetic data. Germline DNA is typically isolated from the white blood cells (buffy coat layer of processed peripheral blood). However, DNA can be extracted from many other biospecimens that could have been collected for any number of clinical or research purposes.[22] Prospective studies often collect peripheral blood at study entry for future PGx analysis. Retrospective studies are sometimes conducted using available samples, genetic data from biobanks, or from a patient's medical record. An alternative approach is to contact potential participants and collect a blood or saliva sample for genetic analysis; however, this approach can be resource intensive.

Besides these critical inclusion criteria, there are certain variables that should be collected to enable exclusion of potential participants from the PGx analysis. These are often based on clinical knowledge, such as excluding patients receiving concomitant medications that may modify the effect of genetics on treatment or excluding patients with comorbid conditions that may modify the effect of treatment on the outcome. Depending on the number of patients who have this confounding variable, it may be possible to adjust for the variable within the PGx analysis instead of excluding the patient.[23] For that reason, and when in doubt, we recommend including the patient and the variable within the data collection and discussing with the study team, including clinical and biostatistical experts, whether to adjust for the variable or exclude the patients.

3. Phenotypes

Introduction

Phenotype is a general term encompassing pharmacological treatment outcomes, such as treatment efficacy or toxicity, or pharmacological characteristics underlying those clinical effects, such as systemic drug concentration. In PGx, phenotype is also used to describe the activity of enzymes and transporters of a patient, as described in the Genetics Section. In this tutorial, phenotype refers to any treatment-

related endpoint that is used as a dependent variable in a PGx association study. Phenotype selection and definition is perhaps the most important and challenging decision in PGx association studies. This section will describe the sources and types of data that can be used as a phenotype, discuss the types of phenotypes that can be used as dependent variables, and provide some recommendations for selecting and defining a phenotype for a PGx association study.

Sources and Types of Data

Within clinical care and research, data are collected in a variety of ways, including objective measurement and subjective assessment. Objective measurements include counts, volumes, masses, concentrations, and durations. Treatment outcomes that are not amenable to measurement fall under subjective assessment and can be assessed by a clinician or the patient.[24] Clinician assessment, either based on clinic notes or documentation within clinical studies, has historically been the primary source of phenotypic data for PGx studies.[25] However, collection of treatment outcomes via patient assessment is increasingly being integrated into clinical care, particularly for subjective toxicities.[26]

These data sources provide a variety of data types, each with their own benefits and drawbacks (**Table 1**). Measurements collected as continuous data can be highly precise and are often most sensitive to the genetic effect, and therefore increase the likelihood of detecting a PGx association.[27] For those reasons, continuous data are often used in discovery-phase PGx research. However, changes in continuous measurements are often not clinically meaningful, so continuous data are commonly translated into ordinal (i.e., improved, no change, worsened) or dichotomous (i.e., change <X% vs. ≥X%) data prior to analysis. Clinician and patient assessments are commonly collected as ordinal (i.e., none, mild, moderate, or severe) data, but can also be translated into dichotomous data to enhance clinical relevance, particularly in validation studies. Finally, for endpoints that are highly dependent on cumulative time or dose, using the time- or dose-at-occurrence as the phenotype will enhance the likelihood of identifying a genetic association,[28, 29] though these cumulative risk models can be somewhat more challenging to translate into clinical practice and may be more relevant to advanced researchers.

As previously described in the Cohort section, phenotypes for PGx association studies are often based on existing data that may have limited availability and accuracy.[15] When accruing a prospective cohort that will be used for PGx analyses, it is critically important to collect accurate phenotypic data, ideally using standardized assessments at pre-specified time points.[30] Regardless of the data source,

abstraction and cleaning of phenotypic data should be conducted while blinded to the genotype data to prevent bias.[31]

Types of Phenotypes

Similar to the types of data, there tends to be a balance between phenotypes that are more sensitive and analytically powerful and others that are more clinically relevant. Highly sensitive phenotypes that are less clinically relevant include surrogate outcomes or quantitative intermediate phenotypes, referred to as endophenotypes. Endophenotypes are often more strongly associated with genetic characteristics since the effect of environmental factors and the number of genes involved is relatively limited.[32] It is therefore sometimes easier to demonstrate the direct genetic effect on an endophenotype. This can be done in smaller discovery-phase studies, followed by determining if genetics are associated with downstream, clinically relevant treatment outcomes in larger validation studies (**Figure 2**). Alternatively, PGx associations with clinical outcomes are often reported first, and endophenotypes can be used to validate the mechanism through which the PGx association is acting.

Pharmacokinetics (PK) is the quintessential PGx endophenotype because it is highly sensitive to genetic variability, specifically in drug metabolizing enzymes and transporters.[33] For drugs with established therapeutic target concentrations, such as tacrolimus,[34] PK can be a clinically relevant surrogate outcome, but for other drugs it is not. Nearly any PK metric can be used as the phenotype in a PGx study, including concentration at a single, clinically relevant timepoint such as a concentration maximum (C_{max}), minimum (C_{min}), or steady-state average ($C_{ss,avg}$). Collection of serial blood samples allows estimation of the full exposure profile by area under the curve (AUC) or clearance (CL), which are particularly relevant for PGx analyses of enzyme and/or transporter activity.[35] Another PK endophenotype that is sometimes used as an indicator of enzymatic activity is the metabolic ratio, which is the ratio of the concentration of the metabolite to the parent compound.[36]

Pharmacodynamic (PD) endophenotypes can also be used within PGx discovery studies.[37] PD endophenotypes include changes in the measurement of a biochemical or physiological marker that are sensitive indicators of treatment response. Changes in International Normalized Ratio (INR) during warfarin treatment[38] is an example of an efficacy PD biomarker, whereas changes in liver function tests[39] to indicate hepatotoxicity is an example of a toxicity PD biomarker. If available, a measurement taken immediately prior to treatment should be used to isolate changes that are attributable to treatment response.

Analyses of PK or PD endophenotypes are useful to establish the direct effect of the genetic variant, but it is typically necessary to confirm the effect on the downstream clinical outcome to justify clinical translation. Relevant clinical outcomes include occurrences of toxicity, particularly severe toxicities that cause permanent morbidity or mortality. On the efficacy side, reductions in disease-related events, such as exacerbations or hospitalizations (or length of stay) and improvements in survival, are clinically relevant outcomes. PGx predictors of these meaningful treatment outcomes are likely to be clinically useful, though it can be very challenging to validate a PGx association since these outcomes are often multi-factorial.[40]

Clinical Outcome Selection and Phenotype Definition

Although any outcome can be used as a phenotype, PGx studies are most likely to yield clinically relevant findings if the phenotype is strongly determined by a single or small set of genes. Clinical outcomes that are strongly PK-dependent are excellent phenotypes because of our relatively robust understanding of the substantial genetic effect on PK.[33] This includes drugs with a narrow therapeutic window and those that require therapeutic drug monitoring. There has also been substantial success identifying useful PGx associations for pro-drugs, which require metabolic activation.[41] Alternatively, clinical outcomes that are completely independent of PK, such as drug-induced hypersensitivity,[42] can be highly dependent on a single genetic variant or gene, leading to profound PGx associations.[43] Other indications that a clinical outcome may have an inherited genetic determinant include those that are non-normally distributed in the population,[42] have large differences in occurrence across races/ethnicities,[44] and treatment outcomes that are similar to an inherited condition.[45] Conversely, multi-factorial clinical outcomes that are partially determined by PK and non-PK factors are more challenging for use as phenotypes for PGx discovery and translation, though there are some successful examples, such as combining *CYP2C9* and *VKORC1* to explain variability in warfarin efficacy.[46]

After selecting a clinical outcome that is potentially genetically determined, the phenotype must be explicitly defined. The phenotype selection and definition should be guided by the putative mechanistic model connecting genetics to the clinical outcome and the objective of the PGx study (**Figure 2**). One potentially useful strategy is to conduct analyses in order of the putative model to confirm each step of the mechanistic pathway. For example, demonstrating that genotype affects PK, and that PK determines the clinical outcome, strongly suggests that genotype will predict the clinical outcome in a sufficiently large validation study.[47] However, if PK does not affect the clinical outcome, then genetic predictors of PK are unlikely to be clinically useful.[40] In addition to considering the mechanistic pathway, phenotype

selection and definition should be consistent with the study objective. Discovery-phase studies may want to use the most sensitive, mechanistically proximal phenotype available, which is often a PK or PD endophenotype. Phenotypes could then get progressively more clinically relevant as the objective moves to replication. Finally, validation studies should select a single, *a priori* defined, clinically relevant endpoint to confirm the genetic association,[48] in preparation for prospective studies and clinical translation.

4. Genotypes

PGx Nomenclature

Understanding the concepts and nomenclature of genetics is vital for PGx investigators. An allele is the genetic base at a given locus, which can be either the more common wild type allele or a less common variant allele. In most cases, humans inherit one allele from each parent and the combination of those two alleles is referred to as their diplotype (**Figure 3**). Allele frequency is the proportion of that allele in the population and is often described in terms of the frequency of the less common allele, or "minor allele frequency (MAF)." Common variants with MAF greater than 5% (or 1%) in the population are referred to as a single nucleotide polymorphisms (SNPs) or single nucleotide variants (SNVs). Many millions of SNPs are cataloged in the NCBI dbSNP Database (https://www.ncbi.nlm.nih.gov/snp/),[49] which includes helpful information such as the SNPs genomic position and MAF in different ethnic groups (**Table 2**).

Alleles are not independently inherited, instead alleles that are nearby are often co-inherited and are said to be in Linkage Disequilibrium (LD).[50] LD exists between nearby SNPs that are inherited in blocks ranging from a few to several hundred kilobases, creating haplotypes of co-occuring SNPs. In addition to SNPs and haplotypes, genetic variations can exist as insertions or deletions of one or more bases, as well as copy number variations (CNVs), where large portions of the genome, including entire genes, are duplicated or deleted.

Candidate Gene/SNP Studies

Candidate Gene Selection

Conventionally, "pharmacogenetics" refers to the investigation of single gene/SNP association with a drug response phenotype, whereas "pharmacogenomics" refers to a genome-wide investigation, described later in this section.[51] The terms are used interchangeably and our use of PGx refers to either. In the candidate gene strategy, researchers evaluate variants within genes with plausible or known

biological mechanisms related to the drug or outcome. Most candidate gene/SNP studies have conventionally focused on genes that are involved in drug PK, especially the enzymes and transporters involved in drug metabolism and disposition. Another common candidate gene is the drug target, which may affect drug response. For instance, *CYP2C9* and *VKORC1* are well-known candidate genes impacting warfarin dosing, as CYP2C9 metabolizes warfarin and VKORC1 is the drug target (**Figure 3**).[46]

Recurrent variants with potential functional consequence in pharmacogenes are assigned * alleles (pronounced "star alleles"). *1 is usually reserved for the more common wild type allele (i.e., *CYP2D6*1*) and is assigned by default when none of the tested variants are detected. Numerical * allele designations (i.e., *CYP2D6*4*) can be defined by a variety of genetic variations including a SNP, haplotype, or a complete gene deletion (e.g., *CYP2D6*5*). To ensure consistency, standardized terminology has been developed to describe the activity of alleles and a patient's predicted activity phenotype for *drug-metabolizing enzymes* and *transporters* (**Table 3**). For some genes, an activity value is assigned to each allele, ranging from 0-1, and the overall activity is assigned by adding the two allele values and translating that sum into a phenotype (i.e., poor (PM), intermediate (IM), normal (NM), rapid (RM), or ultrarapid (UM) metabolizer). These drug metabolizer and transport activity phenotypes are distinct from, and not to be confused with, the phenotype that is the endpoint or dependent variable in the PGx analysis, discussed previously in the Phenotype section. Of note, this phenotype terminology was not always used and some publications use alternative terms such as Extensive Metabolizer (EM), which was replaced by the more intuitive Normal Metabolizer (NM).

It is critical that the process for translating a patient's raw genotype calls to activity phenotypes follows the current best practices, to ensure the validity and interpretability of the study findings. This process is different for each gene and evolves as our understanding of genetics expands, so researchers should review curated information such as that from Clinical Pharmacogenetics Implementation Consortium (CPIC) or PharmGKB (Table 2) when translating genotypes into phenotypes.[52, 53] More advanced investigators may want to use translation software to automate the translation from genotype to phenoytpe.[54] . Ultimately, investigators must decide whether to analyze the PGx association for a single SNP, the combination of variants comprising a predicted activity phenotype, or conduct more extensive genotyping or sequencing to analyze many or all variants within the gene of interest.

Candidate Gene Genotyping

Whenever possible, PGx studies should use high-quality germline DNA that can be easily isolated from whole blood and many healthy tissue types. In some instances, banked samples have been used to

isolate DNA that may have been modified due to disease or sample processing and storage, and this could introduce artifacts that interfere with genotyping accuracy.[55, 56]

There are some specific instances in which blood genotype does not reflect the activity in the tissue of interest. Patients who have had bone marrow transplants will have blood cells from the donor, so the genotypes will not be representative of the subject's other tissues. If patients with bone marrow transplants are included in the study, the germline DNA sample needs to have been collected prior to transplant. Relatedly, patients who have had liver transplants likely have different drug metabolizing enzyme phenotypes than that indicated by blood genotyping. If liver transplant patients are included in studies, donor DNA should be genotyped in addition to the recipient's DNA. Novice investigators should be particularly careful when conducting studies within transplant patients.

A comprehensive review of the many available technologies for genotyping is beyond the scope of this tutorial.[57] The primary consideration when selecting a genotyping technology is the number of variants to genotype. Candidate SNP studies often use single SNP, low throughput genotyping techniques such as TaqMan® or Pyrosequencing. Studies that investigate a group of candidate genes will frequently use genotyping chips or panels. Options include creating a customized panel with the investigator's candidate genes/SNPs (e.g., Assays-by-SEQUENOM [SEQUENOM, San Diego, CA]) or utilizing an existing multi-gene panel. One common approach when conducting a PGx analysis of drug PK is to use a targeted panel array of relevant pharmacogenes such as the DMET™ [Drug Metabolism Enzymes and Transporters], (Affymetrix, Santa Clara, CA),[58] PharmacoScan™ (Thermo Fisher Scientific, Waltham, MA), and VeriDose® Core Panel (Agena Bioscience, San Diego, CA).[59] These off-the-shelf arrays can be highly efficient, though researchers should be careful to select a panel that has adequate coverage of the genes relevant to their drug of interest and variants that are common in the ethnic groups represented within their patient cohort.[4]

Genome Wide Association Studies and Sequencing

An alternative to a candidate gene/SNP approach, typically reserved for more advanced PGx researchers, is to conduct a genome-wide association study (GWAS) (**Table 4**). Genome-wide panels genotype hundreds of thousands of SNPs throughout the genome for simultaneous association testing with treatment outcomes.[60, 61] There has been a rapid increase in GWAS for identifying genetic determinants of a variety of treatment outcomes including efficacy, toxicity, metabolism, and drug-target interactions.[51, 60, 62] GWAS leverages the co-inheritance of SNPs in LD,[63] which enables a single SNP to be used as a marker or "tag" for other SNPs in that haplotype block. By directly genotyping tagSNPs,

investigators can identify genomic regions that harbor causative variants (**Figure 3**). A bioinformatics process known as imputation uses LD to infer the genotypes of SNPs in that region to assist in identifying the likely causative variants.[64-68] Due to differences in LD between ethnic groups, and the predominance of European American ethnicity in genetics reference panels, imputation is less precise for non-European study participants.

GWAS can identify PGx associations for variants and genes that would not have been selected as candidates, which can lead to an improved understanding of the underlying biology of the outcome and/or pharmacology of the drug (e.g. a liver transporter affecting the clearance of a drug that is primarily renally-eliminated).[69] However, due to the huge number of association tests conducted, GWAS studies require significant statistical correction for multiple comparisons, necessitating much larger sample sizes to achieve adequately powered analyses, as described in the statistical methods section.

With continued technological progress, next generation sequencing technologies have resulted in the development of panels for sequencing target genes, the exons of all genes in the genome (whole exome), or whole genomes. Sequencing approaches result in each nucleotide of the genome being sequenced and, therefore, detect every variant in the sequenced region. The cost of sequencing and complexity of data analysis have been barriers to using sequencing to replace GWAS. Further details and discussion of GWAS and sequencing are beyond the scope of this manuscript, but sequencing may be a useful tool for discovery of rare variants with large effects and for follow-up of findings from GWAS.[70] Translating sequencing data to * allele nomenclature is a challenge, particularly given the complexity of properly phasing alleles into haplotypes, but there are tools available to assist advanced PGx researchers with this task.[54, 71]

Genetic Models

Often, the final step in defining the genetic data for PGx analysis is to select a genetic model based on the expected mode of inheritance (i.e., dominant, recessive, or additive). The patient's diplotype is expressed by a combination of two alleles A (major) and a (minor), with possible diplotypes AA (major allele homozygote, i.e., wild type), Aa (heterozygote) and aa (minor allele homozygote). A dominant model would test whether carrying at least one minor (a) allele is associated with the phenotype (i.e., AA vs. (Aa+aa)) whereas a recessive model tests whether the phenotype is associated with carrying two minor alleles (i.e., (AA+Aa) vs. aa). Perhaps the most commonly used genetic model is the additive, or gene-dose, model, which assumes a linear increase in the phenotype with each additional minor a allele (i.e., AA > Aa > aa or AA < Aa < aa). For candidate gene/SNP studies, wherein extensive knowledge of the

gene or SNP and its inheritance is known, researchers may be able to make an informed selection of the appropriate genetic model. On the other hand, studies without this existing knowledge or those testing many SNPs simultaneously typically assume an additive genetic model as it is flexible and has the most power to detect associations, as discussed in the next section.[72-74]

5. Quality Control and Statistical Analysis

Quality Control

As genotyping technologies have improved, the analytical validity of genotype calls (i.e., the accuracy of genotype results) has also improved. Nevertheless, inaccurate genotyping can occur due to assay issues or technical error. In this section, we review several standard approaches to ensuring the quality of genotype data for a PGx study. These approaches are not comprehensive and a plan to ensure genotyping quality control should be developed based on study-specific considerations.

SNP and Sample Call Rate

One of the simplest quality checks for genotype data are using the proportion of SNPs or samples that were successfully genotyped, referred to as "call rate". Studies often remove poor performing SNPs or samples based on inadequate call rates.[23, 75] The removal is based on the assumption that assays or samples with missing genotype calls are also likely to have incorrect genotype calls. The SNP call rate is the number of samples successfully genotyped divided by the total number of samples for which genotyping was attempted. There is no universal SNP call rate cutoff, but studies commonly use a cutoff of 95%-99%, below which, SNPs are removed from the analysis. Similarly, the sample call rate is the number of SNPs successfully genotyped in that sample divided by the total number of SNPs for which genotyping was attempted. Removing low quality SNPs and samples based on call rates improves the genotype data quality and minimizes the potential effect of genotyping errors on the study results.

Hardy Weinberg Equilibrium

Hardy Weinberg Equilibrium (HWE) is a mathematical expression of the expected distribution of alleles and genotypes in a population under certain conditions, such as a lack of natural selection and lack of genetic drift. Similar to its use in population genetics, investigators can use this equation to test for evidence of genotyping error.[23, 75] If genotypes for a SNP do not follow the expected HWE distribution, a possible explanation is that the SNPs have been genotyped incorrectly. HWE testing can identify excess heterozygosity, a term used for when there is an excess number of individuals with the Aa genotype. This can result from a SNP assay that cannot effectively distinguish between alleles (i.e., A vs. a) or from

inadvertent mixture of two different DNA samples. HWE can be tested by comparing the actual distribution of alleles with the distribution expected based on the known MAF, or the actual MAF within the sample, using chi square or Fisher's exact test. The p-value threshold below which HWE deviation is considered unacceptable is often corrected for multiple comparisons. SNPs observed to be below the threshold are removed from the final analysis under the assumption that genotyping error was observed. However, another common explanation for HWE departure is racial admixture, since genotype frequencies can vary substantially based on ancestry.[76] Therefore, in diverse cohorts, HWE should be tested within each racial/ethnic group. Another important consideration for HWE testing is that patients with a specific disease may be enriched for certain SNPs, causing departure from HWE. Removal of these SNPs may actually remove the SNPs with the most relevant effects on the phenotype of interest. The test for HWE in case control studies is often conducted in controls alone rather than in the whole population, based on the assumption that enrichment for SNPs is less likely to occur in the control population. In any event, testing for HWE can be an effective tool to ensure genotyping quality, but should be thoughtfully applied to studies with particular consideration for cohort selection.

Considerations for Statistical Analysis

As with any biomedical study, statistical approaches for PGx association testing will be guided by the study design with particular consideration for the nature of the phenotypic data (i.e., continuous, ordinal, or dichotomous, normal or non-normal distribution; paired versus un-paired) and whether there is a need to account for confounding variables **(Figure 4)**. Although a comprehensive review of statistical approaches is beyond the scope of this tutorial, this section focuses on statistical considerations that are particularly relevant to PGx association testing, with a major focus on minimization of false positive and negative findings, as appropriate for the study objective.

Multiple Comparisons Adjustment

PGx studies often include more than one SNP of interest, which increases the likelihood of a false positive result due to multiple comparisons.[23] Assuming a standard alpha (α) =0.05 (i.e., $p<0.05$), on average one out of every twenty tested associations will be statistically significant by chance alone. False positive results can be minimized by using a more stringent alpha. The most common correction for multiple comparisons is the Bonferroni correction, which divides the alpha by the number of independent association tests conducted. This becomes particularly important in studies with large numbers of SNPs, such as GWAS.[23, 75] A GWAS including a million SNPs would on average detect 50,000 significant SNP associations by chance using $\alpha=0.05$. Regardless of the number of SNPs in the GWAS,

genome-wide significance is typically set at $\alpha=5x10^{-8}$ (alpha=0.05/1,000,000),[77] as illustrated by the horizontal line in the standard visualization of GWAS results using a Manhattan Plot in Figure 4. This threshold may be overly conservative, since SNPs are in LD and their associations are not completely independent, and alternative thresholds could be considered prior to analysis.[78, 79] One commonly used less-conservative alternative approach is the False Discovery Rate (FDR) adjustment.[80]

Testing associations for one SNP with multiple phenotypes, or with one phenotype using multiple genetic models (i.e., dominant, recessive, and additive), also increases the number of association tests and risk of false positives if proper statistical correction is not applied. As the number of association tests increases, the corrected α decreases and statistical significance becomes more difficult to achieve, increasing risk of false negatives. For this reason, it is advised to limit the number of association tests conducted. As in other scientific fields, researchers will often specify a primary hypothesis that includes a single genotype and phenotype of primary interest and conduct that analysis with an uncorrected α=0.05. All other analyses are then considered exploratory, also conducted using an uncorrected α=0.05, and reported as hypothesis-generating.

Statistical Power

Power is the ability of the study to observe a true PGx association, thus avoiding false negatives. As discussed above, multiple comparisons correction decreases α and makes significance more difficult to achieve, thus reducing study power and increasing false negative risk. Analytical power is determined during study design and influenced by the statistical test, sample size, phenotype variability, and the magnitude of genetic effect on the phenotype.[81] Power increases with larger cohorts and with more patients within each genotype group. As such, association tests get more powerful as the SNP MAF increases, meaning that studies of rare SNPs are often underpowered, even with large cohorts. Publicly available MAF estimates[82] can be used to conduct power determinations prior to initiating a project. Another determinant of the size of genotype groups is the selected genetic model, described in the prior section.[83] Studies with potential power concerns should avoid recessive genetic models (AA+Aa vs. aa), unless there is compelling prior knowledge to justify their use, since the homozygous variant (aa) group is the smallest and these analyses have limited power. In terms of phenotypic variability, power increases with higher variability of continuous endpoints and with higher event rate of dichotomous endpoints. Finally, power increases substantially as the magnitude of the effect of the genotype on the phenotype, which is referred to as the effect size, increases.[84, 85] Effect size can be expressed as a difference between means of a continuous endpoint or differences in event rates of a dichotomous

endpoint. Detailed instructions for how to conduct a sample size determination is beyond the scope of this article and should typically be the responsibility of the study biostatistician, though relatively simple sample size calculators are freely available online.[86, 87]

Population Stratification

Admixture of groups with different ethnic background, termed population stratification, was previously mentioned as a potential cause of HWE departure. Population stratification can also result in false positive PGx associations due to confounding of the risk of toxicity and uneven genotype distribution across ethnic groups.[23, 75] For example, because individuals from Southeast Asia have a higher rate of carbamazepine-induced Stevens-Johnson Syndrome (SJS), a GWAS of this phenotype would identify significant associations for any SNP that is differentially carried in Asians vs. non-Asian patients, regardless of whether these SNPs actually cause SJS.[88] These false associations can be avoided by conducting analyses within a single race/ethnicity group or adjusting for self-reported race/ethnicity or genetically-defined ancestry. In GWAS, it is standard practice for advanced researchers to derive principal components that mathematically describe the racial contribution to genetic variation and adjust for these components in the statistical analyses to reduce the potential for population stratification.

Replication and Validation

Similar to other branches of science, many reported SNP associations cannot be successfully reproduced.[6, 7] These false positive findings are typically from discovery studies that test many potential associations without proper statistical correction for multiple comparisons. False positives must be differentiated from actual, robust associations that may be useful in clinical practice.[5] Researchers can try to assess the robustness of the associations detected in their study by conducting sensitivity analyses, in which slight changes are made to the patient inclusion/exclusion, genetic model, or phenotype definition. Genetic association for which these minor adjustments dramatically affect the effect size and p-value may be false positives or may be insufficiently robust for clinical translation.

While sensitivity analyses can be informative, replication and validation are the optimal methods for differentiating valid PGx associations from false positives, as described in the Study Objective section.[23, 89, 90] Replication and validation of a clinical PGx association is often challenging due to the relative infrequency of having access to large patient cohorts who have been exposed to a specific drug, meet study eligibility criteria, and have available phenotypic data. In cases where no such cohort exists, researchers often turn to cohorts that are as similar as possible, such as cohorts of patients treated with

a drug in the same class and/or having similar phenotypic data. Another possible approach to collect data supporting a clinical PGx association is to perform pre-clinical (i.e., in cells, tissues, or animals) validation experiments or to use publicly available datasets, such as GTEx and Haploreg, to generate further evidence of a SNP's influence on gene function and expression.[91, 92] While not a formal replication, results from these datasets can bolster the evidence that the SNP has a causal effect on the phenotype.

Prediction Accuracy

The accuracy of the genotype to predict the phenotype is a critical determinant of the potential clinical utility of a PGx association.[93] For dichotomous outcomes, multiple statistical tools are available to assess the predictive power of a PGx variant, including the area under the receiver operating characteristic (AUROC) curve and the calculation of specificity, sensitivity, positive predictive value (PPV), and negative predictive value (NPV). Working with the prior example of carbamazepine-induced SJS, *HLA-B*15:02* is predictive of this adverse drug reaction.[88] In the Han Chinese population, *HLA-B*15:02* testing has estimated sensitivity=98% and specificity=97%.[94] This means that a positive *HLA-B*15:02* test detects 98% of individuals that will have carbamazepine-induced SJS and a negative test detects 97% of individuals that will not have carbamazepine-induced SJS. Similarly, the estimated PPV is 7.7% and the NPV is 100%, meaning that only 7.7% of those with positive *HLA-B*15:02* test will have carbamazepine-induced SJS, but 100% of those with a negative *HLA-B*15:02* test will not have carbamazepine-induced SJS. These results have important implications for the clinical utility of the test. For instance, the low PPV means that many patients who test positive, and do not receive carbamazepine, would not have developed SJS if given carbamazepine. However, given the availability of similarly effective alternative antiepileptic agents, pharmacogenetic-based antiepileptic treatment may still be clinical useful.

Conclusion

This tutorial has described critical considerations when performing a PGx association study, starting with determining the research objective, selecting the cohort, defining a phenotype that is consistent with the objective, genotyping via candidate or genome-wide approaches, and planning an appropriate statistical analysis. Thinking through these major decisions when developing a PGx association study will maximize the chances of success for novice investigators. Although beyond the scope of this tutorial, comprehensive reporting of the methods and results of PGx association studies in peer-reviewed manuscripts is critical. We strongly recommend that novice researchers review prior publications describing best practices for reporting PGx studies,[95] including the recently published STROPS

guidelines.[96] PGx association studies can be powerful tools for discovery, replication, or validation of associations between inherited genetic variation and treatment outcomes, providing the evidence necessary for future clinical translation to improve efficacy and reduce toxicity of pharmacotherapy.

Acknowledgements

The authors would like to thank Ethan Poweleit, Zachary Taylor, Sarah Board, Shana Littleton, and Ana Lopez Medina and for a thoughtful review of the manuscript. Figures were created with Biorender.com.

Data Type	Description	Benefits	Drawbacks	Examples	Ideal PGx Study Types
Continuous	Values	• Most sensitive to genetic variation • Highest analytical power • Established statistical methods • Precise estimate of genetic effect	• Sensitive to non-genetic effects • Often multifactorial • Least clinically useful	• Drug concentration • Change in blood pressure	Discovery and Replication
Ordinal	≥3 ordered groups	• Somewhat sensitive to genetic variation • Maintains useful ordering • Sometimes clinically relevant	• Statistical methods not as well established • Not easily clinically translated	• Severity (none, mild, moderate) • Grade (0, 1, 2, 3) • Continuous data with ≥2 cut points	Replication
Dichotomous	Two groups	• Most clinically relevant • Most often used (case/control) • Established statistical methods • Most clinically translatable	• Least sensitive for analysis • Least genetically dependent	• Yes/No event occurred • Case/control classification • Continuous data with 1 cut point • Ordinal data with 1 cut point	Validation
Time- (dose-) to event	Dichotomous, but accounts for time (or dose)	• Similar benefits as dichotomous • Accounts for time or dose • Censor patients who drop out	• Similar drawbacks as dichotomous • Clinical translation is difficult	• Survival time • Cumulative dose at toxicity	Any

Table 1: Data Types and Attributes for Phenotypes

Table 2: Helpful Resources for PGx Investigators

Resource	Web Link	Features/Attributes
Pharmacogene Variation Consortium	https://www.pharmvar.org/	Catalogue allelic variation of genes including the SNP or SNPs in the haplotype and their resulting functional activity
Clinical Pharmacogenetics Implementation Consortium (CPIC)	www.cpicpgx.org	Reports variant frequencies in many ethnic cohorts. Also provides expert consensus recommendations for genotype to phenotype translation and publishes clinical practice guidelines for validated gene/drug pairs that are indexed in PubMed
The Pharmacogenomics Knowledgebase (PharmGKB)	www.pharmgkb.org	Leading worldwide resource for PGx knowledge, allowing searches by drug, gene, or SNP, ultimately directing the end-user to freely accessible, evidence-graded primary PGx literature
Findbase	https://findbase.org/#/	Online resource cataloguing frequencies of clinically relevant pharmacogenomic biomarkers in various populations
ClinGen	https://www.clinicalgenome.org/	Repository for clinically relevant genes and variants including pharmacogenomic variants
NCBI dbSNP Database	https://www.ncbi.nlm.nih.gov/snp/	Public database of known SNPs including genomic position and minor allele frequency across cohorts

Table 3: Standardized Terms for Allelic and Phenotypic Activity[97]

Activity	Allele function (all genes)	Drug metabolizing enzyme phenotypes	Transporter phenotypes
Highest ⬇ Lowest	Increased function	Ultrarapid metabolizer	Increased function
	Increased function	Rapid metabolizer	Increased function
	Normal function	Normal metabolizer	Normal function
	Decreased function	Intermediate metabolizer	Decreased function
	No function	Poor metabolizer	Poor function
Unknown	Unknown/uncertain function		

Table 4: Considerations for Selecting Candidate or Genome-wide Pharmacogenetic Study

	Candidate SNP/gene Study	Genome-wide association Study
Study Objective	Best for replication and validation	Best for discovery
Approximate Cohort Size	100s	1000s
Number of Genetic Variants	1-1000s	1,000,000+
Gene Selection	Genes related to plausible mechanisms such as enzymes and transporters for pharmacokinetic associations, or prior associations	Selection not required (all genes analyzed simultaneously)
Variant Selection	Known functionally consequential variants or prior associations	tagSNPs that are informative of nearby variants within haplotype blocks
Typical Genetic Model	Selected based on prior knowledge or reported association	Additive
Visualization of Association	Phenotype stratified by genotype using bar, box, or survival plots	P-value of association for each variant using Manhattan Plot
Critical advantage	Less statistical correction to detect associations	Identify associations outside of candidate genes, efficiency.

Critical Limitation	Only detect associations for variants selected as candidates	Requires large cohorts, statistical correction, and is costly

REFERENCES

1. Chidambaram AG, Josephson M. Clinical research study designs: The essentials. *Pediatr Investig.* 2019;3:245-252. doi: 210.1002/ped1004.12166. eCollection 12019 Dec.

2. Teutsch SM, Bradley LA, Palomaki GE, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. *Genetics in medicine : official journal of the American College of Medical Genetics.* 2009;11:3-14.

3. Luzum JA, Pakyz RE, Elsey AR, et al. The Pharmacogenomics Research Network Translational Pharmacogenetics Program: Outcomes and Metrics of Pharmacogenetic Implementations Across Diverse Healthcare Systems. *Clin Pharmacol Ther.* 2017;102:502-510. doi: 510.1002/cpt.1630. Epub 2017 Jun 1009.

4. Arwood MJ, Chumnumwat S, Cavallari LH, Nutescu EA, Duarte JD. Implementing Pharmacogenomics at Your Institution: Establishment and Overcoming Implementation Challenges. *Clin Transl Sci.* 2016;9:233-245.

5. Xiao R, Boehnke M. Quantifying and correcting for the winner's curse in genetic association studies. *Genetic epidemiology.* 2009;33:453-462.

6. Ioannidis JP. To replicate or not to replicate: the case of pharmacogenetic studies: Have pharmacogenomics failed, or do they just need larger-scale evidence and more replication? *Circ Cardiovasc Genet.* 2013;6:413-418; discussion 418.

7. Ioannidis JP. Why most published research findings are false. *PLoS medicine.* 2005;2:e124.

8. Safran C, Bloomrosen M, Hammond WE, et al. Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. *J Am Med Inform Assoc.* 2007;14:1-9. doi: 10.1197/jamia.M2273. Epub 2006 Oct 1131.

9. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA.* 2000;284:3043-3045.

10. McGuire AL, Majumder MA. Two cheers for GINA? *Genome Med.* 2009;1:6. doi: 10.1186/gm1186.

11. Chevrier R, Foufi V, Gaudet-Blavignac C, Robert A, Lovis C. Use and Understanding of Anonymization and De-Identification in the Biomedical Literature: Scoping Review. *J Med Internet Res.* 2019;21:e13484. doi: 13410.12196/13484.

12. Pulley J, Clayton E, Bernard GR, Roden DM, Masys DR. Principles of human subjects protections applied in an opt-out, de-identified biobank. *Clin Transl Sci.* 2010;3:42-48. doi: 10.1111/j.1752-8062.2010.00175.x.

13. Fullerton SM, Lee SS. Secondary uses and the governance of de-identified data: lessons from the human genome diversity panel. *BMC Med Ethics.* 2011;12:16.:10.1186/1472-6939-1112-1116.

14. Wang X, Kattan MW. Cohort Studies: Design, Analysis, and Reporting. *Chest.* 2020;158:S72-S78. doi: 10.1016/j.chest.2020.1003.1014.

15. Conway M, Berg RL, Carrell D, et al. Analyzing the heterogeneity and complexity of Electronic Health Record oriented phenotyping algorithms. *AMIA Annu Symp Proc.* 2011;2011:274-283. Epub 2011 Oct 2022.

16. DiPietro NA. Methods in epidemiology: observational study designs. *Pharmacotherapy.* 2010;30:973-984. doi: 910.1592/phco.1530.1510.1973.

17. Ratain MJ, Miller AA, McLeod HL, Venook AP, Egorin MJ, Schilsky RL. The cancer and leukemia group B pharmacology and experimental therapeutics committee: a historical perspective. *Clin Cancer Res.* 2006;12:3612s-3616s. doi: 3610.1158/1078-0432.CCR-3606-9008.

18. Wright AF, Carothers AD, Campbell H. Gene-environment interactions--the BioBank UK study. *Pharmacogenomics J.* 2002;2:75-82. doi: 10.1038/sj.tpj.6500085.

19. Porzsolt F, Wiedemann F, Becker SI, Rhoads CJ. Inclusion and exclusion criteria and the problem of describing homogeneity of study populations in clinical trials. *BMJ Evid Based Med.* 2019;24:92-94. doi: 10.1136/bmjebm-2018-111115. Epub 112018 Dec 111119.

20. Haga SB, LaPointe NM. The potential impact of pharmacogenetic testing on medication adherence. *Pharmacogenomics J.* 2013;13:481-483. doi: 410.1038/tpj.2013.1033. Epub 2013 Sep 1033.

21. Sterne JAC, Hernán MA, McAleenan A, Reeves BC, Higgins JPT. Assessing risk of bias in a non-randomized study. *Cochrane Handbook for Systematic Reviews of Interventions.* 2019:621-641.

22. Caudle KE, Keeling NJ, Klein TE, Whirl-Carrillo M, Pratt VM, Hoffman JM. Standardization can accelerate the adoption of pharmacogenomics: current status and the path forward. *Pharmacogenomics.* 2018;19:847-860. doi: 810.2217/pgs-2018-0028. Epub 2018 Jun 2219.

23. Grady BJ, Ritchie MD. Statistical Optimization of Pharmacogenomics Association Studies: Key Considerations from Study Design to Analysis. *Curr Pharmacogenomics Person Med.* 2011;9:41-66. doi: 10.2174/187569211794728805.

24. Basch E, Iasonos A, McDonough T, et al. Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study. *Lancet Oncol.* 2006;7:903-909.

25. Hertz DL. Concerns regarding use of patient-reported outcomes in biomarker studies of chemotherapy-induced peripheral neuropathy. *Pharmacogenomics J.* 2019;19:411-416.

26. Basch E, Reeve BB, Mitchell SA, et al. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *J Natl Cancer Inst.* 2014;106:dju244. doi: 210.1093/jnci/dju1244. Print 2014 Sep.

27. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ.* 2006;332:1080.

28. Prentice RL, Kalbfleisch JD, Peterson AV, Jr., Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics.* 1978;34:541-554.

29. Mega JL, Close SL, Wiviott SD, et al. Cytochrome P-450 Polymorphisms and Response to Clopidogrel. *New England Journal of Medicine.* 2009;360:354-362.

30. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Insitute; 2010.

31. Ahern TP. Pharmacoepidemiology in Pharmacogenetics. *Adv Pharmacol.* 2018;83:109-130.:10.1016/bs.apha.2018.1004.1001. Epub 2018 May 1010.

32. Maranville JC, Cox NJ. Pharmacogenomic variants have larger effect sizes than genetic variants associated with other dichotomous complex traits. *Pharmacogenomics J.* 2016;16:388-392.

33. Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. *N Engl J Med.* 2011;364:1144-1153.

34. Birdwell KA, Decker B, Barbarino JM, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for CYP3A5 Genotype and Tacrolimus Dosing. *Clin Pharmacol Ther.* 2015;98:19-24. doi: 10.1002/cpt.1113. Epub 2015 Jun 1003.

35. Yukawa E. Population-based investigations of drug relative clearance using nonlinear mixed-effect modelling from information generated during the routine clinical care of patients. *J Clin Pharm Ther.* 1999;24:103-113. doi: 110.1046/j.1365-2710.1999.00204.x.

36. Inaba T, Vinks A, Otton SV, Kalow W. Comparative pharmacogenetics of sparteine and debrisoquine. *Clin Pharmacol Ther.* 1983;33:394-399. doi: 310.1038/clpt.1983.1051.

37. Hertz DL, McLeod HL. Using pharmacogene polymorphism panels to detect germline pharmacodynamic markers in oncology. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2014;20:2530-2540.

38. International Warfarin Pharmacogenetics C, Klein TE, Altman RB, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. *The New England journal of medicine.* 2009;360:753-764.

39. Xu CF, Johnson T, Wang X, et al. HLA-B 57:01 Confers Susceptibility to Pazopanib-Associated Liver Injury in Patients with Cancer. *Clin Cancer Res.* 2016;22:1371-1377. doi: 1310.1158/1078-0432.CCR-1315-2044. Epub 2015 Nov 1376.

40. Hertz DL, Rae JM. One step at a time: CYP2D6 guided tamoxifen treatment awaits convincing evidence of clinical validity. *Pharmacogenomics.* 2016;17:823-826.

41. Begg EJ, Helsby NA, Jensen BP. Pharmacogenetics of drug-metabolizing enzymes: the prodrug hypothesis. *Pharmacogenomics.* 2012;13:83-89.

42. Daly AK. Using genome-wide association studies to identify genes important in serious adverse drug reactions. *Annual Review of Pharmacology and Toxicology.* 2012;52:21-35.

43. Martin MA, Kroetz DL. Abacavir pharmacogenetics--from initial reports to standard of care. *Pharmacotherapy.* 2013;33:765-775.

44. Chung WH, Hung SI, Hong HS, et al. Medical genetics: a marker for Stevens-Johnson syndrome. *Nature.* 2004;428:486.

45. Chua KC, Kroetz DL. Genetic advances uncover mechanisms of chemotherapy-induced peripheral neuropathy. *Clin Pharmacol Ther.* 2017;101:450-452.

46. Johnson JA, Gong L, Whirl-Carrillo M, et al. Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. *Clinical pharmacology and therapeutics.* 2011;90:625-629.

47. Hertz DL. Exploring pharmacogenetics of paclitaxel- and docetaxel-induced peripheral neuropathy by evaluating the direct pharmacogenetic-pharmacokinetic and pharmacokinetic-neuropathy relationships. *Expert Opin Drug Metab Toxicol.* 2021;6:1-13.

48. Institute of Medicine: Evolution of Translational omics: Lessons Learned and the Path Forward. Washington, DC: TheNational Academies Press; 2012.

49. Sherry ST, Ward M, Sirotkin K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res.* 1999;9:677-679.

50. Elston RC, Satagopan JM, Sun S. Genetic terminology. *Methods Mol Biol.* 2012;850:1-9.:10.1007/1978-1001-61779-61555-61778_61771.

51. Pirmohamed M. Pharmacogenetics and pharmacogenomics. *Br J Clin Pharmacol.* 2001;52:345-347.

52. Barbarino JM, Whirl-Carrillo M, Altman RB, Klein TE. PharmGKB: A worldwide resource for pharmacogenomic information. *Wiley Interdiscip Rev Syst Biol Med.* 2018;10:e1417. doi: 1410.1002/wsbm.1417. Epub 2018 Feb 1423.

53. Relling MV, Klein TE, Gammal RS, Whirl-Carrillo M, Hoffman JM, Caudle KE. The Clinical Pharmacogenetics Implementation Consortium: 10 Years Later. *Clin Pharmacol Ther.* 2020;107:171-175. doi: 110.1002/cpt.1651. Epub 2019 Nov 1005.

54. Sangkuhl K, Whirl-Carrillo M, Whaley RM, et al. Pharmacogenomics Clinical Annotation Tool (PharmCAT). *Clin Pharmacol Ther.* 2020;107:203-210. doi: 210.1002/cpt.1568. Epub 2019 Sep 1017.

55. Goetz MP, Sun JX, Suman VJ, et al. Loss of Heterozygosity at the CYP2D6 Locus in Breast Cancer: Implications for Germline Pharmacogenetic Studies. *Journal of the National Cancer Institute.* 2014;107:10.1093/jnci/dju1401.

56. Hertz DL, Kidwell KM, Douglas JA, et al. DNA derived from archival tumor specimens can be used for germline pharmacogenetic analyses. *Pharmacogenomics.* 2020;21:899-902. doi: 810.2217/pgs-2020-0071. Epub 2020 Jul 2229.

57. Aquilante CL, Zineh I, Beitelshees AL, Langaee TY. Common laboratory methods in pharmacogenomics studies. *Am J Health Syst Pharm.* 2006;63:2101-2110.

58. Arbitrio M, Di Martino MT, Scionti F, et al. DMET™ (Drug Metabolism Enzymes and Transporters): a pharmacogenomic platform for precision medicine. *Oncotarget.* 2016;7:54028-54050.

59. Lemieux Perreault LP, Zaïd N, Cameron M, Mongrain I, Dubé MP. Pharmacogenetic content of commercial genome-wide genotyping arrays. *Pharmacogenomics.* 2018;19:1159-1167. doi: 1110.2217/pgs-2017-0129. Epub 2018 Oct 1151.

60. Motsinger-Reif AA, Jorgenson E, Relling MV, et al. Genome-wide association studies in pharmacogenomics: successes and lessons. *Pharmacogenet Genomics.* 2013;23:383-394.

61. di Iulio J, Rotger M. Pharmacogenomics: what is next? *Front Pharmacol.* 2011;2:86.:10.3389/fphar.2011.00086. eCollection 02011.

62. Crowley JJ, Sullivan PF, McLeod HL. Pharmacogenomic genome-wide association studies: lessons learned thus far. *Pharmacogenomics.* 2009;10:161-163.

63. Meng Z, Zaykin DV, Xu CF, Wagner M, Ehm MG. Selection of genetic markers for association analyses, using linkage disequilibrium and haplotypes. *Am J Hum Genet.* 2003;73:115-130. doi: 110.1086/376561. Epub 372003 Jun 376565.

64. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet.* 2010;11:499-511.

65. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 2009;5:e1000529.

66. Vergara C, Parker MM, Franco L, et al. Genotype imputation performance of three reference panels using African ancestry individuals. *Hum Genet.* 2018;137:281-292.

67. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 2009;5:e1000529. doi: 1000510.1001371/journal.pgen.1000529. Epub 1002009 Jun 1000519.

68. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol.* 2010;34:816-834. doi: 810.1002/gepi.20533.

69. Treviño LR, Shimasaki N, Yang W, et al. Germline Genetic Variation in an Organic Anion Transporter Polypeptide Associated With Methotrexate Pharmacokinetics and Clinical Effects. *Journal of Clinical Oncology.* 2009;27:5972-5978.

70. Gulilat M, Lamb T, Teft WA, et al. Targeted next generation sequencing as a tool for precision medicine. *BMC Medical Genomics.* 2019;12:81.

71. Lee SB, Wheeler MM, Patterson K, et al. Stargazer: a software tool for calling star alleles from next-generation sequencing data using CYP2D6 as a model. *Genet Med.* 2019;21:361-372. doi: 310.1038/s41436-41018-40054-41430. Epub 42018 Jun 41436.

72. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. *PLoS Comput Biol.* 2012;8:e1002822.

73. Jorgensen AL, Williamson PR. Methodological quality of pharmacogenetic studies: issues of concern. *Stat Med.* 2008;27:6547-6569.

74. Horita N, Kaneko T. Genetic model selection for a case-control study and a meta-analysis. *Meta Gene.* 2015;5:1-8.

75. Srinivasan BS, Chen J, Cheng C, et al. Methods for analysis in pharmacogenomics: lessons from the Pharmacogenetics Research Network Analysis Group. *Pharmacogenomics.* 2009;10:243-251.

76. Deng HW, Chen WM, Recker RR. Population admixture: detection by Hardy-Weinberg test and its quantitative effects on linkage-disequilibrium methods for localizing genes underlying complex traits. *Genetics.* 2001;157:885-897.

77. Panagiotou OA, Ioannidis JP. What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *Int J Epidemiol.* 2012;41:273-286.

78. Wilkening S, Chen B, Bermejo JL, Canzian F. Is there still a need for candidate gene approaches in the era of genome-wide association studies? *Genomics.* 2009;93:415-419.

79. Duncan LE, Ostacher M, Ballon J. How genome-wide association studies (GWAS) made traditional candidate gene studies obsolete. *Neuropsychopharmacology.* 2019;44:1518-1523.

80. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med.* 1990;9:811-818. doi: 810.1002/sim.4780090710.

81. Attia J, Ioannidis JP, Thakkinstian A, et al. How to use an article about genetic association: A: Background concepts. *JAMA.* 2009;301:74-81. doi: 10.1001/jama.2008.1901.

82. The International HapMap Project. *Nature.* 2003;426:789-796.

83. Attia J, Ioannidis JP, Thakkinstian A, et al. How to use an article about genetic association: B: Are the results of the study valid? *JAMA.* 2009;301:191-197. doi: 110.1001/jama.2008.1946.

84. Ross S, Anand SS, Joseph P, Paré G. Promises and challenges of pharmacogenetics: an overview of study design, methodological and statistical issues. *JRSM Cardiovasc Dis.* 2012;1:cvd.2012.012001. doi: 012010.011258/cvd.012012.012001.

85. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461:747-753. doi: 710.1038/nature08494.

86. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G Power 3.1: tests for correlation and regression analyses. *Behav Res Methods.* 2009;41:1149-1160. doi: 1110.3758/BRM.1141.1144.1149.

87. Dupont WD, Plummer WD, Jr. Power and sample size calculations for studies involving linear regression. *Control Clin Trials.* 1998;19:589-601. doi: 510.1016/s0197-2456(1098)00037-00033.

88. Leckband SG, Kelsoe JR, Dunnenberger HM, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for HLA-B genotype and carbamazepine dosing. *Clin Pharmacol Ther.* 2013;94:324-328. doi: 310.1038/clpt.2013.1103. Epub 2013 May 1021.

89. Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature.* 2007;447:655-660. doi: 610.1038/447655a.

90. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet.* 2005;6:95-108. doi: 110.1038/nrg1521.

91. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 2012;40:D930-934. doi: 910.1093/nar/gkr1917. Epub 2011 Nov 1097.

92. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* 2013;45:580-585. doi: 510.1038/ng.2653.

93. Attia J, Ioannidis JP, Thakkinstian A, et al. How to use an article about genetic association: C: What are the results and will they help me in caring for my patients? *JAMA.* 2009;301:304-308. doi: 310.1001/jama.2008.1993.

94. Hung SI, Chung WH, Jee SH, et al. Genetic susceptibility to carbamazepine-induced cutaneous adverse drug reactions. *Pharmacogenet Genomics.* 2006;16:297-306. doi: 210.1097/1001.fpc.0000199500.0000146842.0000199504a.

95. McDonagh E, Whirl-Carrillo M, Altman RB, Klein TE. Enabling the curation of your pharmacogenetic study. *Clin Pharmacol Ther.* 2015;97:116-119. doi: 110.1002/cpt.1015. Epub 2014 Dec 1020.

96. Chaplin M, Kirkham JJ, Dwan K, Sloan DJ, Davies G, Jorgensen AL. STrengthening the Reporting Of Pharmacogenetic Studies: Development of the STROPS guideline. *PLoS Med.* 2020;17:e1003344. doi: 1003310.1001371/journal.pmed.1003344. eCollection 1002020 Sep.

97. Caudle KE, Sangkuhl K, Whirl-Carrillo M, et al. Standardizing CYP2D6 Genotype to Phenotype Translation: Consensus Recommendations from the Clinical Pharmacogenetics Implementation Consortium and Dutch Pharmacogenetics Working Group. *Clin Transl Sci.* 2020;13:116-124.

Figure Legends

Figure 1: An overview of important considerations when planning and conducting a pharmacogenetic association study. Abbreviations: SNP, single nucleotide polymorphism; PGx, pharmacogenetic(s); GWAS, genome-wide association study.

Figure 2: Mechanistic chain of Pharmacogenetics Associations. It may be easiest to identify a pharmacogenetic (PGx) effect on a proximal surrogate, such as the effect of pharmacokinetic pharmacogenetics (PK PGx) on a PK surrogate of drug concentrations. A similar process can be used to test for pharmacodynamic pharmacogenetics (PD PGx) on a PD surrogate of biochemical response. If genetics affects one of these surrogate endophenotypes, it can then be tested for an effect on a clinically relevant treatment outcome in validation studies in preparation for potential clinical translation.

Figure 3. A, Inheritance pattern of alleles to create haplotypes. A pedigree chart is drawn in the middle with metabolizer status indicated inside the shape (NM, normal metabolizer; IM, intermediate metabolizer). Each person's alleles for the CYP2D6 gene are shown under their picture, boxes indicate exons, darker boxes indicate coding exons with a variant, X indicates splicing variant. The CYP2D6*4 allele has a splice site variant and an activity value (AV) of 0. The CYP2D6*10 allele contains two variants, in exons 1 and 9, conferring an activity value of 0.25. The mother's diplotype is *1/*4, with an activity value of 1, which corresponds with an intermediate metabolizer phenotype. The father's diplotype is *1/*10, with an activity value of 1.25, which corresponds with an normal metabolizer phenotype. The baby's diplotype is *1/*10, with an activity value of 1, which corresponds with an intermediate metabolizer phenotype. The boy's diplotype is *4/*10, with an activity value of 0.25, which corresponds with an intermediate metabolizer phenotype.
B. Selection of candidate single nucleotide polymorphisms (SNPs) for warfarin dose phenotype and illustration of the differences between variants captured from sequencing, genome-wide association study (GWAS), and candidate SNP genotyping for functional SNPs.
Figure 4: Examples of visual representations for different types of data acquired during a PGx association study. Continuous phenotype data such as for drug clearance is often summarized in box plots by patient genotype. Ordinal and dichotomous phenotype data are often represented by proportions of patients by genotype in histograms. Time to event data is plotted by genotype in Kaplan-Meier plots which summarize the proportions of patients at risk for an event at a given time point after study enrollment. Genome-wide association study data is generally represented in a Manhattan plot,

which plots the chromosomal location of SNP variation along the x axis and the -log(p value) for each SNP along the y axis.

1 Study objective

Discovery
Testing many SNP-phenotype associations

Replication
Retesting several SNP-phenotype associations in a separate population

Validation
Retesting single SNP-phenotype associations in a separate population

Translation
Prospectively testing PGx-guided treatment

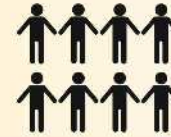
Implementation
Using PGx to inform treatment decisions

2 Study cohort

Diverse

vs.

Homogenous



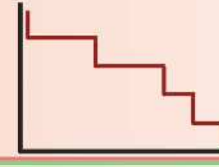
3 Phenotype

Exposure

Adverse Reaction

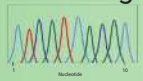
Response

Survival



4 Genotype

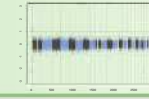
Single SNP



Candidate gene

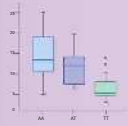
Targeted panel

GWAS



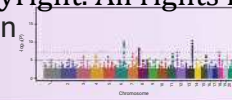
5 Analysis

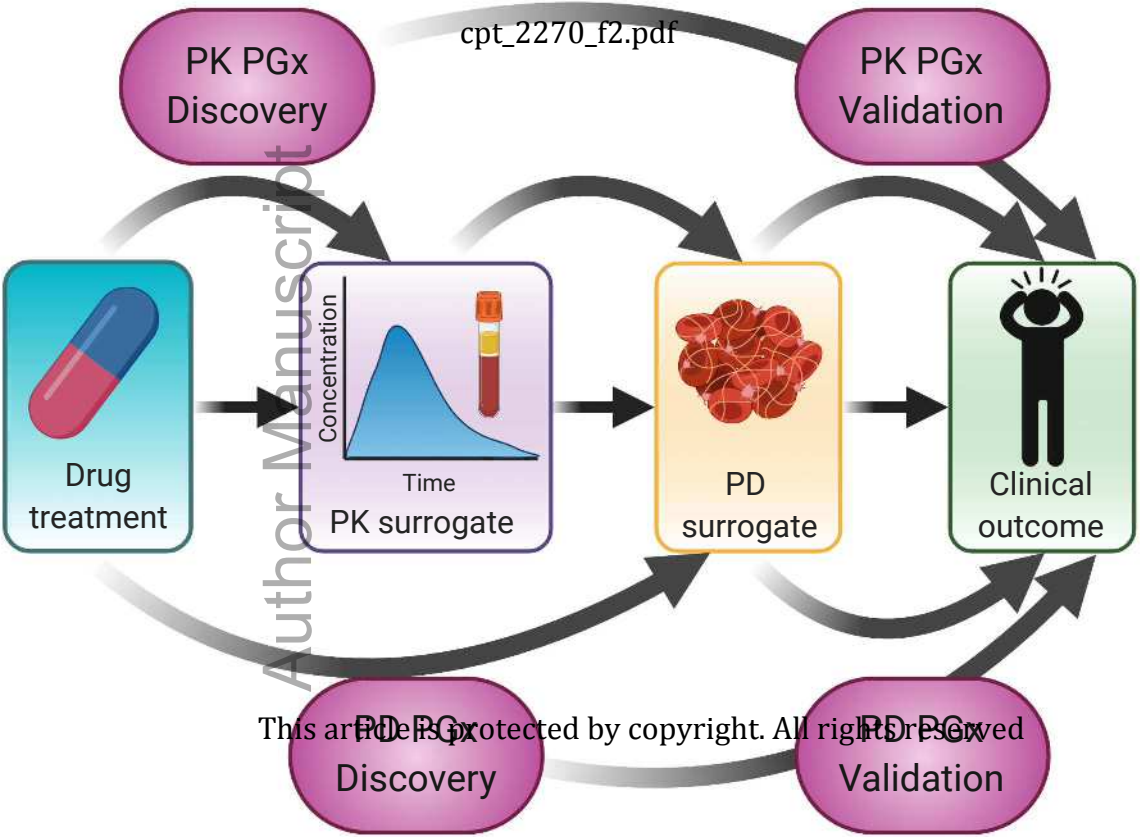
Quality Control

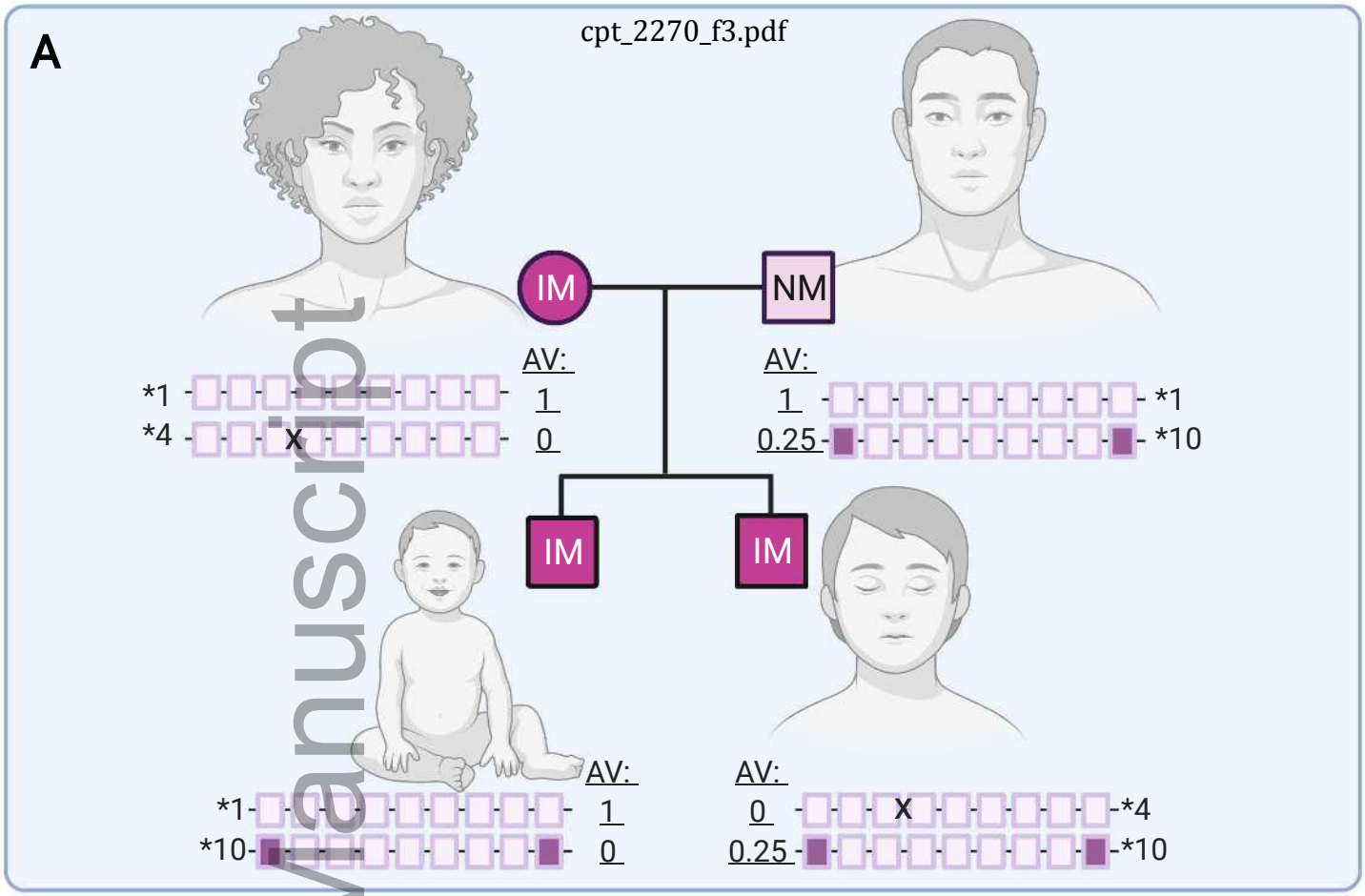


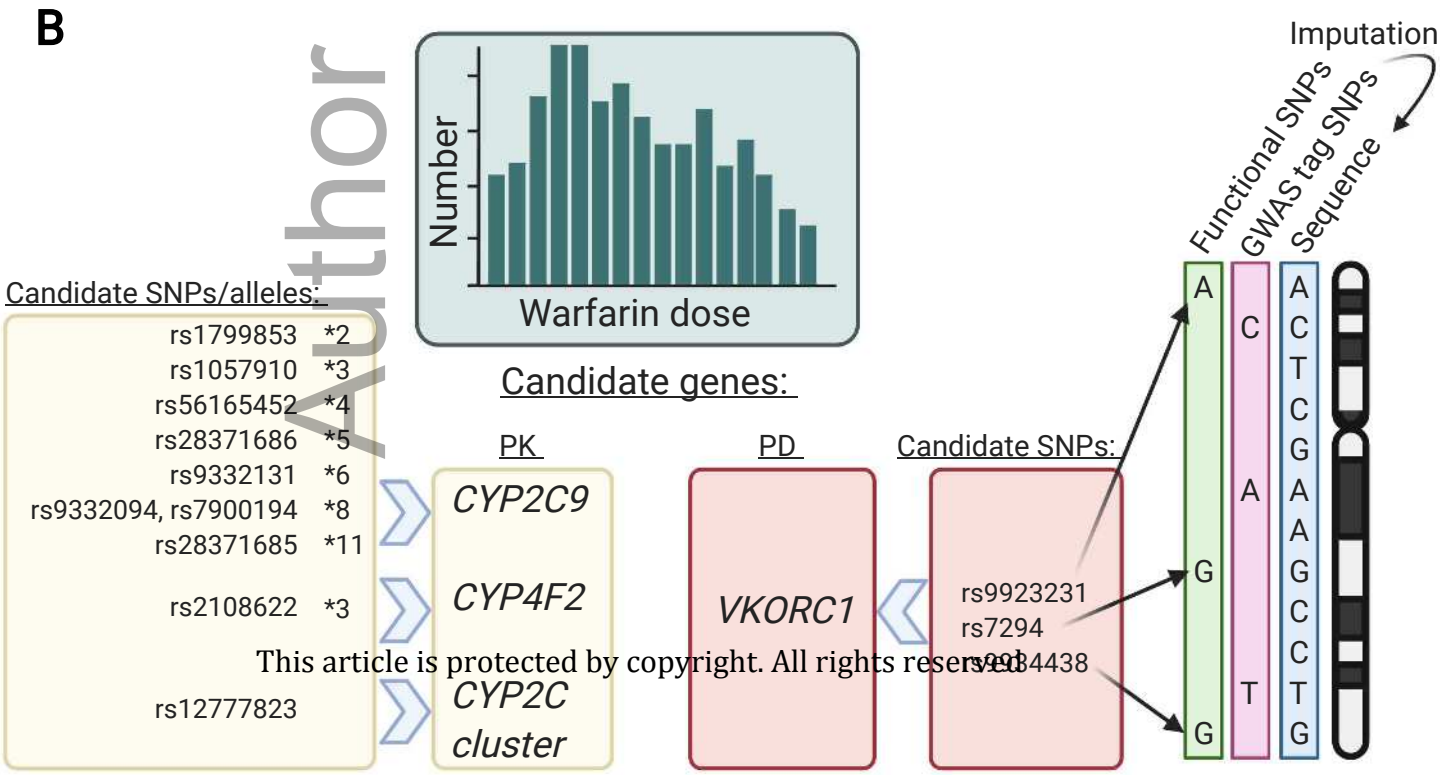
Statistical tests

Multiple testing correction

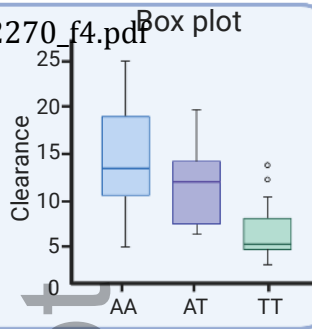




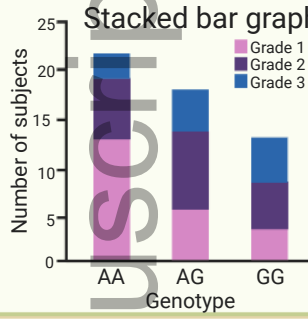




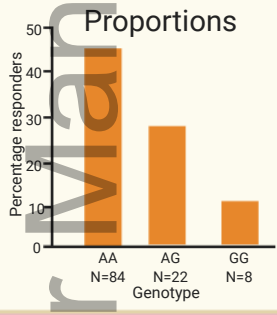
Continuous data



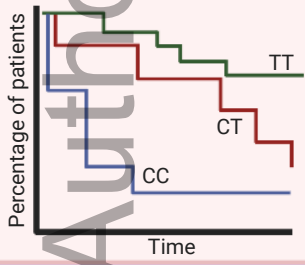
Ordinal data



Dichotomous data



Time to event data



GWAS data