Clinical Evidence Supporting Pharmacogenomic Biomarker Testing Provided in US Food and Drug Administration Drug Labels

ABSTRACT

Background: Genetic biomarkers that predict a drug's efficacy or likelihood of toxicity are assuming increasingly important roles in the personalization of pharmacotherapy, but there is concern that there may be insufficient evidence linking use of some biomarkers to clinical benefit. Nevertheless, information about the use of biomarkers appears in the drug label for many prescription medications. This may add confusion to the clinical decision-making process.

Objective: To evaluate the evidence supporting pharmacogenomic biomarker testing in drug labels, how frequently testing is recommended, and completeness of citation of the supporting studies.

Methods: We used publicly-available databases from the US Food and Drug Administration (FDA) to identify drug labels that described the use of a biomarker and evaluated whether the label contained or referenced convincing evidence of its clinical validity (*i.e.*, the ability to predict phenotype) and clinical utility (*i.e.*, the ability to improve clinical outcomes) using guidelines published by the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) group. We graded the completeness of the citation of supporting studies and determined whether the label recommended incorporation of biomarker test results in therapeutic decision-making.

Results: Of the 119 drug-biomarker combinations, only 36% (n=43) had labels that provided convincing clinical validity evidence while 15% (n=18) provided convincing evidence for clinical utility. Fifty-one percent of the labels (51%, n=61) made recommendations based on the results of a biomarker test; 30% of these contained

convincing clinical utility data. A full description of supporting studies was included in 11% (n=13) of the labels.

Discussion: Fewer than one-sixth of drug labels contained or referenced convincing evidence for clinical utility of biomarker testing while over half made recommendations based on biomarker test results. It may be premature to include biomarker testing recommendations in drug labels when convincing data linking testing to patient outcomes does not exist.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express deep gratitude to my Principal Investigator, Dr. Niteesh K. Choudhry, for his invaluable mentorship and constant offering of keen insights throughout the course of this scholarly pursuit, from study conception to publication.

Much appreciation also to Mr. William J. Canestaro, fellow researcher and co-author, for his contributions to the study's methodology, statistical analysis, and data interpretation, all while residing 3,000 miles away in Seattle.

And, of course, thank you to all the incredible individuals in my research division – the Division of Pharmacoepidemiology and Pharmacoeconomics at Brigham and Women's Hospital – for mentoring and supporting me through the numerous research endeavors I have carried out with the group since coming up to Boston. Their work to promote the betterment of public health and well-being is nothing short of awe-inspiring, and, after working with the division for more than 4 years, I can attest to the fact that the brilliance of their minds is surpassed only by their generosity and kindness. To Dr. Jerry Avorn, the Division Chief – thank you for welcoming a green, untested pharmacy student with no relevant research background into your research division in the spring of 2011, and for wholeheartedly supporting my decision to conduct a research year there after my second year in medical school.

The findings of this scholarly project have been previously published under the same title in *JAMA Internal Medicine* in December 2014.

GLOSSARY

FDA	– US Food and Drug Administration
CYP450	– Cytochrome P450
INR	– International normalized ratio
EGAPP	– Evaluation of Genomic Applications in Practice and Prevention
SSRI	– Selective serotonin reuptake inhibitor

INTRODUCTION

Background

The ability to target and tailor drug therapies based on genetic information has created much hope that personalization of pharmacotherapy will revolutionize health care. This enthusiasm is supported by a vast literature evaluating a variety of biomarkers with polymorphisms that have predictive significance and can potentially improve the efficacy or safety profiles of drugs treating numerous disease states, including cancer, depression, and cardiovascular disease.[1 2 3 4] Based on this, the US Food and Drug Administration (FDA) has included pharmacogenomic information in the physician prescribing information ("drug labels") of more than 100 FDA-approved drugs. For example, the label for the oral anticoagulant warfarin suggests dosage adjustments based on a patient's genotype for *CYP2C9* and *VKORC1*,[5] while the label for abacavir, an antiretroviral, recommends avoiding its use in patients who screen positive for the *HLA-B*5701* allele.[6] Frueh and colleagues' study in 2008 noted that almost two-thirds of the labels containing information on human biomarkers referred to polymorphisms in cytochrome P450 (CYP450) metabolism and that 1 in 4 patients filled prescriptions for at least one drug whose label contained human biomarker information.[7] The FDA maintains a continuously updated list of these drugs, along with their associated biomarkers, in the FDA Table of Pharmacogenomic Biomarkers in Drug Labeling.[8]

Biomarker information in drug labels

Prescription drug labels are an important source of information about drug therapies for many providers[9] and the information contained in them also appears in other frequently consulted references such as UpToDate[10], Micromedex[11], and the Physicians' Desk Reference.[12] Both a drug's initial label and subsequently revised labels are reviewed and regulated by the FDA, with initiatives requiring additional data and formatting changes being implemented periodically to facilitate accurate and efficient dissemination of information to prescribers and, ultimately, patients.[13]

Recently, Stanek and colleagues conducted a national survey to assess physicians' perceptions of pharmacogenomics testing as well as their willingness to utilize these tools to guide medication therapy. Among the over 10,000 physicians who responded, almost all

agreed that patients' genetic profiles may affect their response to pharmacotherapy, but only 1 in 10 felt adequately informed about pharmacogenomic testing[14]; in addition, roughly half reported consulting FDA-approved drug labels for information regarding pharmacogenomic testing. Given how inadequately informed physicians feel about the appropriate use of pharmacogenomic biomarkers, the information and recommendations included in drug labels should provide clear guidance by not only being evidence-based but also directly relevant to clinical decision-making.

However, despite their inclusion in drug labels, the use of many biomarkers do not appear to be clearly associated with health benefits. For example, although the label for warfarin contains a recommended dosing algorithm based on *CYP2C9* and *VKORC1* polymorphisms, the "clinical utility" of genotype-based dosing for this medication (*i.e.*, the ability to improve clinical outcomes with biomarker testing compared to management without genetic testing[15]) remains unclear. Three different randomized controlled trials exploring the clinical utility of genotype-guided warfarin were simultaneously published in the *New England Journal of Medicine* in 2013. Kimmel and colleagues[16] reported no differences in anticoagulation control between the genotype-guided and clinically guided groups at the end of 4 weeks of therapy, in terms of both percentage of time spent in therapeutic INR range as well as rates of major clinical outcomes, including international normalized ratio (INR) of 4 or greater, major bleeding, or thromboembolism. Verhoef and colleagues[17] similarly found no changes in percentage of time in the therapeutic INR range or incidence of bleeding and thromboembolic events between the genotype-guided and clinically guided groups at the end of the 12-week study period. By contrast, Pirmohamed and colleagues[18] found that the genotype-guided patient group achieved a greater mean percentage of time in therapeutic INR range along with significantly fewer incidences of INR measurements of 4 or higher at the end of the 12-week study period. A meta-analysis published the following year that included these three studies along with six other randomized clinical trials which compared genotype-guided and clinical dosing of warfarin in adults found no differences between the two groups in both the percentage of time spent in the therapeutic INR range or risk of major complications.[19]

Furthermore, inclusion of potentially tenuous recommendations or associations within a drug's label may encourage clinicians to order tests or lead them to change

therapies based on limited evidence, even if the label does not recommend explicit action based on biomarker status. The potentially suboptimal patient outcomes that may result from such actions is compounded by the financial implications, with the cost associated with genetic tests projected to increase at more than twice the rate of overall health spending and reach $10 billion in the US by 2015.[20 21]

Evaluation of Genomic Applications in Practice and Prevention Working Group

To help address this evidence gap, the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group was established in 2005 by the Centers for Disease Control and Prevention to develop a systematic method of evidence-based assessment of pharmacogenomic tests and other genomic applications.[22] The group focused their evaluative process on three previously-described components of genetic testing: analytic validity, clinical validity, and clinical utility. A genetic test's analytic validity is defined by EGAPP as its "ability to accurately and reliably measure the genotype (or analyte) of interest in the clinical laboratory, and in specimens representative of the population of interest." Clinical validity is defined by EGAPP as a test's "ability to accurately and reliably predict the clinically defined disorder or phenotype of interest." Clinical validity encompasses analytic validity, and thus establishment of the latter component is a prerequisite for reliable further evaluation of the former. EGAPP defines a genetic test's clinical utility as "the evidence of improved measurable clinical outcomes, and its usefulness and added value to patient management decision-making compared with current management without genetic testing"; a test's clinical utility relates most directly to the level of patient benefit it offers.

To allow for the systematic grading of these three components (analytic validity, clinical validity, and clinical utility) for a given genetic test, EGAPP established a hierarchy of study designs and internal validity rubrics, tailored to each component.[22] In general, well-designed systematic reviews or meta-analyses and large randomized controlled trials or cohort studies were ranked as higher quality evidence for a particular component, while consensus guidelines, unpublished and/or non-peer reviewed studies, and expert opinions were ranked lowest. Using this hierarchy of study designs, the data for each component was then rated as convincing, adequate, or inadequate to reflect the degree to which the

observed effect (either of net health benefit or lack thereof) is likely to be an accurate assessment rather than due to flaws in the study's methodology.

In addition to creating this framework for evaluating genetic tests, EGAPP also utilized this methodology to conduct its own systematic reviews of evidence supporting the adoption of pharmacogenomic testing in clinical practice. However, as of May 2015, the group has only been able to conduct three systematic reviews investigating the effect of pharmacogenomic testing on drug therapy – CYP450s and selective serotonin reuptake inhibitors (SSRIs)[23]; *UGT1A1* and irinotecan[24]; and *EGFR* and cetuximab and panitumumab.[25] EGAPP concluded that there was not enough good-quality data to evaluate the effect of CYP450 testing on SSRIs, and found adequate analytic validity and clinical validity and inadequate clinical utility for the effect of *UGT1A1* testing and irinotecan. For the association between *EGFR* testing and cetuximab and panitumumab, the group concluded that there was inadequate clinical validity and no studies to evaluate clinical utility.

Furthermore, while it is important that drug labels present evidence-based data directly relevant to clinical decision-making, these documents should also convey this information in a succinct and self-sufficient manner, as well as include complete citations for referenced studies to allow prescribers to efficiently locate the full studies for further independent assessment of the evidence. This is especially important given the increasingly time-constrained environment that physicians and other prescribers are forced to navigate, fueled at least in part by exponential growths in scientific knowledge and clinical practice guidelines as well as ever-increasing administrative responsibilities, all working to decrease the amount and quality of time clinicians are able to spend with their patients and in the conductance of clinical care.[26] [27] [28]

AIMS OF STUDY

We sought to apply EGAPP's method of evidence-based assessment to evaluate the strength of clinical evidence supporting the clinical validity and clinical utility of biomarker testing in all drug labels containing pharmacogenomic biomarker information. We also sought to assess how frequently biomarker testing is recommended in these drug labels, as well as examine how completely the supporting studies are cited in the labels.

METHODS

Data sources

Identification of drugs and biomarkers

We identified medications that contained biomarker information in their labels from the FDA Table of Pharmacogenomic Biomarkers in Drug Labeling. This database contains the drug name, therapeutic area, biomarker, and section(s) in the drug label that contains mention of the biomarker. For drugs with more than one associated biomarker, we evaluated each drug-biomarker relationship separately. We accessed this database on April 18, 2013, and utilized the most recently updated list available at that time for our study.

Gathering of drug labels

For each medication, we gathered the earliest available drug label containing mention of its associated biomarker(s) from Drugs@FDA, another publicly-accessible FDA database available through the agency's website that lists regulatory actions, including initial approvals and drug labeling changes.[29] Labels were found on the "Label and Approval History" page for each drug. We analyzed each accessible label from the earliest available one following the drug's initial approval onward until we identified a label containing mention of the drug's associated biomarker; we used this label for our study. We focused on the earliest available label containing mention of the associated biomarker in order to evaluate the quality of the cited evidence supporting testing recommendations as well as accessibility of these supporting studies at the time when physicians could have first become aware of this pharmacogenomic information through the drug label. We could not gather the drug labels from this FDA database for four of the drugs in our study - nefazodone, propafenone, protriptyline, thioridazine - and instead gathered their labels from either the PharmGKB[30] (nefazodone) or the National Institutes of Health database[31] (propafenone, protriptyline, thioridazine). Because information about a single biomarker may have been included in the label of several different drugs and because several drug labels contained information about more than one biomarker, the drug-biomarker combination was our unit of analysis.

Data extraction and evaluation

Two of the authors (BW and WJC) independently evaluated the labels for the quality of cited clinical evidence, the completeness of the citation of supporting studies, and the presence of recommendations for the incorporation of biomarker testing into clinical practice.

The authors first constructed a preliminary rating framework that was applied to a random sample of 12 (10%) of drug labels. The rating framework was then refined to resolved initial disagreements. Using the updated framework, these two authors independently reviewed each label to determine the level of evidence it presented to support the use of biomarker testing in clinical practice; the presence of biomarker testing recommendations; and the completeness of citation of supporting studies. Remaining disagreements were resolved by discussion and consensus among all investigators.

Strength of evidence for clinical validity and utility

Using guidelines from the EGAPP group, we scored the robustness of evidence supporting use of biomarker testing in clinical practice on two dimensions: clinical validity and clinical utility. To mirror how practicing clinicians are likely to interpret the evidence cited within a label, we graded the quality of evidence supporting a biomarker's clinical validity and utility based on the information presented within the drug labels and evaluated other relevant studies only when a citation was included in the label allowing for their location on the PubMed/MEDLINE database.[32] We did not gather or evaluate other sources of information physicians may utilize to inform their practice. Our rationale for doing this was two-fold: drug labels should be self-sufficient in presenting time-constrained prescribers with the evidence base and references supporting incorporation of biomarker testing into clinical practice, and the content presented in the labels themselves is often used to inform other widely used tertiary-level drug information resources.

Clinical validity

Clinical validity of a genetic test is defined by the EGAPP group as "its ability to accurately and reliably predict the clinically defined disorder or phenotype of interest."[22]

We graded the quality of the evidence supporting a biomarker test's clinical validity for its associated drug based on the robustness of the studies that had been carried out at the time of the inclusion of pharmacogenomic information in a drug's label and which were either described or adequately cited in the label to allow for their location in the PubMed/MEDLINE database. We used the criteria presented in **Table 1**, adapted from EGAPP, to rate the evidence for clinical validity as convincing, adequate, or incomplete (the "incomplete" grade is based on the criteria that EGAPP uses to define data as "inadequate").

If the drug label contained evidence addressing the clinical validity of a drug-biomarker relationship, including pharmacokinetic studies, but provided an insufficient study description and citation to identify its design, we rated the evidence for clinical validity as "incomplete." For example, we considered a label for Drug X that listed "clinical studies" as having demonstrated a decreased hemoglobin level in patients with G6PD deficiency but which did not provide or cite information to identify the study design as "incomplete" evidence for clinical validity of the Drug X-G6PD relationship. We also rated the evidence for clinical validity as "incomplete" if the drug label did not contain or cite any evidence that addressed the clinical validity of the drug-biomarker relationship.

We assumed that all study results listed in the FDA label were published and peer-reviewed and erred on the side of considering evidence as convincing when determining whether the evidence demonstrated the ability to "accurately and reliably predict the clinically defined disorder or phenotype of interest" as well as when evaluating (using the classification framework in **Table 1**) the quality of a particular study design (*e.g.*, "high quality longitudinal cohort study") or its consistency/homogeneity (*e.g.*, "systematic review/meta-analysis of well-designed longitudinal cohort studies with homogeneity"). We did not require reporting of statistical testing to classify evidence as convincing for clinical validity. Also, the absolute risks and risk differences did not need to be reported or cited in the label to achieve adequate or convincing clinical validity classification.

The following examples of hypothetical study results in a drug label demonstrate our approach to rating clinical validity evidence:

Example 1: *"A randomized controlled study demonstrated that poor metabolizers of Biomarker X have about a 50% increase in Drug X exposure and about a 25% decrease in exposure to the active metabolite compared to extensive metabolizers."*
We would rate the evidence presented in example 1 as "convincing" for clinical validity because it cites a high quality study design to demonstrate that individuals with varying levels of the biomarker have different metabolic profiles for the associated medication.

Example 2: *"A pharmacogenomic analysis of 100 patients evaluated polymorphisms of Biomarker X and its potential association with hyperbilirubinemia during Drug X treatment. In this study, genotype A was associated with a significant increase in the risk of hyperbilirubinemia relative to genotype B or genotype C."*
We would rate the evidence presented in example 2 as "adequate" for clinical validity because it cites a "case control study with good reference standards" (**Table 1**) to demonstrate that individuals with different levels of biomarker activity manifest varying risks of developing hyperbilirubinemia when taking this drug.

Example 3: *"Poor metabolizers of Biomarker X have about a 50% increase in Drug X exposure and about a 25% decrease in exposure to the active metabolite compared to extensive metabolizers."*
This example presented pharmacokinetic studies addressing the clinical validity of the drug-biomarker relationship. Since insufficient description was provided to identify the study design, we would rate the evidence as "incomplete."

Example 4: *"There appears to be heterogeneity in the activity level of Biomarker X among different individuals."*
Because there is no mention of evidence supporting the clinical validity of the drug-biomarker relationship, we would rate the evidence presented as "incomplete."

Clinical utility

Clinical utility of a genetic test is defined by the EGAPP group as the "evidence of improved measurable clinical outcomes, and its usefulness and added value to patient

management decision-making compared with current management without genetic testing."[22] As such, in order for evidence in drug labels to support the clinical utility of biomarker testing, we required that it demonstrated improved patient outcomes when biomarker testing was incorporated into treatment decisions compared to when no testing was performed. For targeted therapies in which the medication was developed specifically to interfere with or manipulate a certain biomarker to have its intended effect (*e.g.*, trastuzumab was developed specifically to target the *Her2/neu* biomarker), we did not require the presence of a non-testing group for the evidence to support clinical utility but instead focused our assessment on studies that evaluated the drug's efficacy in the intended patient population (*i.e.*, those who screened positive for the targeted biomarker).

As with clinical validity, we graded the quality of evidence supporting a biomarker test's clinical utility for its associated drug based on the robustness of the studies carried out which were either described in the label or adequately cited to allow for their location in the PubMed/MEDLINE database (**Table 1**). As with validity, we erred on the side of considering evidence as convincing when determining whether the evidence demonstrated "improved measurable clinical outcomes" and when evaluating the quality of a particular study design. We also did not require reporting of statistical testing to classify evidence as convincing for clinical utility. However, the pertinent absolute efficacy/risks and risk differences did need to be reported or cited in the label to achieve adequate or convincing clinical utility classification. Finally, either hard clinical endpoints (*e.g.*, mortality) or robust surrogate endpoints (*e.g.*, time within therapeutic INR) were sufficient to establish clinical utility. The following examples demonstrate our approach:

Example 1: *"A randomized, placebo-controlled trial demonstrated that Drug X is 32% less effective in achieving local tumor control in patients with Disease X who lack Biomarker X."* We would rate the evidence presented in example 1 as "convincing" evidence supporting clinical utility because the evidence is consistent with EGAPP's definition of clinical utility (through the study's demonstration that individuals with varying levels of the biomarker have altered clinical outcomes) and the type of study conducted to generate this evidence falls in the "convincing" category ("at least one large randomized controlled trial"). In addition, the study's pertinent efficacy/risk results were reported.

Example 2: *"Drug X is less effective in patients with Disease X who lack Biomarker X."*
Although the evidence presented in example 2 is consistent with the definition of clinical utility, the label did not adequately describe the trial that produced this conclusion for us to be able to identify the study design, nor did it present the pertinent efficacy/risk results. Therefore, we would rate it as "incomplete."

Example 3: *"A randomized-controlled study demonstrated that poor metabolizers of Biomarker X have about a 50% increase in Drug X exposure and about a 25% decrease in exposure to the active metabolite compared to extensive metabolizers."*
The example presented only demonstrates clinical validity. If there is no mention of the drug-biomarker relationship elsewhere in the label that is consistent with EGAPP's definition of clinical utility, we would rate the evidence presented as "incomplete."

Example 4: *"A randomized-controlled trial showed that there is no relevant effect of genetic variation in Biomarker X on Drug X's effect on blood pressure, with poor metabolizers and extensive metabolizers experiencing an average decrease in systolic blood pressure of 12 mm Hg and 13 mm Hg, respectively (p=0.38)."*

In this example, we judged the finding of a lack of clinical effect based on genetic status to be consistent with the demonstration of clinical utility because this evidence can be useful for physicians who are deciding among various pharmacotherapeutic options for a patient's medical needs, especially if clinical utility evidence is not available for the other options or the evidence for these other therapeutics demonstrates significant alterations in clinical effect based on the patient's genetic status. We would therefore grade the clinical utility evidence for this particular example as "convincing."

Completeness of citation of supporting studies

We graded the completeness of citation of supporting studies in drug labels as full, partial, or none. Although drug labels should not be expected to contain a full description of all evidence supporting the clinical validity and clinical utility of incorporating biomarker

testing into clinical decision making, we believe that all relevant studies should be fully cited or referenced so that clinicians can easily gather the full studies for further review and assessment. A grade of "full" was given to labels that included a sufficient citation of supporting studies in a separate references section or mention of the study's name or other identifier (*e.g.*, *PREDICT-1* study for abacavir) such that the studies can be located in the PubMed/MEDLINE database. A grade of "partial" was given to drug-biomarkers whose labels that did not identify the study by name/identifier or in a separate references section but which described the study design and/or results. For clinical validity, these results include variations in pharmacokinetic levels of a drug based on status of the associated biomarker. A grade of "none" was reserved for labels that made no description of the supporting studies. When the completeness of description of supporting studies in drug labels differed between clinical validity and clinical utility, we used the higher grade for our analysis.

Treatment recommendations contained within the label

We evaluated whether the label recommended incorporation of the biomarker test result in therapeutic decision-making. These recommendations were categorized as being based directly on a drug's mechanism of action, as indicated by mention in either the *Indications* or *Mechanism of Action* sections of the drug label (*e.g.*, trastuzumab's targeting of *Her2/neu* overexpression and busulfan's treatment of chronic myelogenous leukemia, a condition characterized by a translocation in the Philadelphia chromosome biomarker, were interpreted as implicit recommendations for biomarker testing before treatment initiation), being based on drug-biomarker associations (dosage adjustment, contraindication/avoidance, and follow-up lab testing), or being absent. A recommendation was considered to be present even when it advocated for alterations in therapy of another drug taken concurrently as a result of drug-drug interactions, as long as the interaction was due to polymorphisms in the given biomarker of the associated drug. For example, a recommendation would be considered present in the label of Drug X if the following statement was present: *"In healthy volunteers who were extensive metabolizers of Biomarker X, Drug X 20mg daily was given in combination with 20mg Drug Y every 12 hours. This resulted in increases in steady state Drug Y AUC values...dosage adjustment of Drug Y may be*

necessary and it is recommended that Drug Y be initiated at a reduced dose when it is given with Drug X."

Statistical analysis

We used descriptive statistics to summarize the supporting evidence and recommendations contained in the drug label for each drug-biomarker combination. We then performed a prespecified Fisher's exact test to determine whether biomarkers for targeted therapies (*i.e.*, medications developed specifically to interfere with or manipulate a biomarker to have its intended effect) demonstrated a different proportion of convincing data supporting clinical utility compared to non-targeted therapies, as well as to compare the quality of cited evidence in labels with and without biomarker testing recommendations. Since numerous cancer agents were specifically developed to be targeted therapies, whereas most neuropsychiatric drug-biomarker combinations are discovered after approval and are related to adverse events, we also performed a prespecified Fisher's exact test to determine whether biomarkers for oncology and neuropsychiatry drugs demonstrated different proportions of convincing evidence supporting clinical utility in their labels compared to other biomarkers.

RESULTS

We identified 119 drug-biomarker combinations representing 107 drugs and 39 unique biomarkers (**Table 2**). The majority of these drug-biomarkers (75, 63%) are intended to reduce the occurrence of adverse drug events, while the remainder (44, 37%) relate to the drugs' efficacy. The most common therapeutic areas covered by these biomarkers were oncology (37, 31%), neuropsychiatry (33, 28%), gastroenterology (9, 8%), infectious disease (9, 8%), and cardiovascular disease (8, 7%) (**Table 3**).

Evidence of clinical validity and utility

Thirty-six percent of drug labels (n=43) provided convincing evidence of the clinical validity of the biomarker (*i.e.*, the ability of the biomarker to predict the phenotype of interest), while 15% (n=18) provided convincing evidence that the use of the biomarker has clinical utility (*i.e.*, the biomarker's ability to improve clinical outcomes).

Table 3 contains examples of our grading approach. In the case of abacavir, for instance, the association between an increased incidence of hypersensitivity reactions and patients with the *HLA-B*5701* allele was supported by both convincing clinical validity and clinical utility evidence via a large randomized, double-blinded study comparing pre-therapy screening for the allele with no pre-therapy screening; the pre-screened group experienced a hypersensitivity reaction of 3.4 percent compared with 7.8 percent for the control group. On the other hand, the association between fluorouracil and an increased risk of adverse events in patients with *DPD* enzyme deficiency was supported by only one case report of life-threatening systemic toxicity in a patient with deficiency for this enzyme, translating to incomplete clinical validity and utility. See **Appendix Tables** for individual biomarker evaluations. Sixty-four percent of drug labels (n=76) did not provide convincing evidence for either clinical validity or utility. Biomarkers for cancer drugs were much more likely to demonstrate convincing evidence supporting clinical utility in their labels compared to all other biomarkers (14/37 [38%] vs. 4/82 [5%], $p<0.001$), while neuropsychiatry biomarkers were less likely to demonstrate convincing clinical utility evidence in their labels than the remaining biomarkers (0/33 [0%] vs. 18/86 [21%], $p<0.001$). Targeted therapies consisted mainly of oncology drugs (26/34, 76%). Fifty percent of targeted therapies (n=17) contained convincing data supporting clinical utility in their labels compared with 1% (1/85) of non-targeted therapies ($p<0.001$). Abacavir, whose label recommended against its use in patients who screened positive for the *HLA-B*5701* allele, was the only non-targeted therapy whose biomarker was supported by convincing clinical utility evidence.

Eleven percent (n=13) of labels contained a full citation of supporting studies, while 30% (n=36) made no mention of the scientific literature supporting the biomarker it discussed.

Treatment recommendations contained within the label

Fifty-one percent (n=61) of the labels made recommendations about how clinical decisions should be based on the results of a biomarker test, with 29% (n=34) being based directly on the drug's mechanism of action and 23% (n=27) being based on drug-biomarker associations. Among biomarkers with neither convincing clinical utility nor

validity data, 32% (24/76) of labels still contained testing recommendations. The drug label for the psychotropic drug iloperidone, for instance, recommended consideration of dose adjustments based on patients' *CYP2D6* status, despite lack of both clinical utility and validity data. Similarly, among labels making recommendations, only 30% (18/61) provided convincing clinical utility data. Labels that made testing recommendations were more likely to provide convincing clinical utility data compared to labels that made no recommendations (30% vs. 0%, $p<0.001$).

DISCUSSION

Overall, our analysis revealed deficiencies in the evidence provided in drug labels supporting the use of many pharmacogenomic biomarkers, with fewer than one-sixth of labels containing or citing convincing evidence for clinical utility and almost two-thirds even lacking convincing data for clinical validity. This deficiency was especially prominent among biomarkers of non-targeted therapies, which constituted more than 70% of our study sample.

The limited amount of convincing evidence for the clinical utility of pharmacogenomic biomarkers in the labels we reviewed is perhaps not surprising given the difficulty of demonstrating that a test actually alters clinical outcomes rather than simply predicting a disorder or phenotype. Nevertheless, the primary goal and potential of pharmacogenomics and personalized medicine is to change patient outcomes. In our opinion, it is premature to include testing recommendations in labels when such utility data is neither described nor cited in this resource, even if it exists elsewhere. It could be argued that high-quality data showing that individuals with certain polymorphisms metabolize a drug differently may be relevant to patient care and therefore should be included in drug labels as the basis for biomarker testing. However, other than predictive value, the inclusion of this type of information does not provide guidance as to what providers should actually do when they get the test results. The case of *CYP2C19* testing for clopidogrel illustrates this well. Polymorphisms of this gene are associated with poorer drug metabolism and a higher risk of thrombotic events, prompting testing recommendations for these variants to be added to the drug's label in 2010. Yet, the consequences of the clinical actions that might rationally be taken based on this

information have still not been convincingly evaluated. If this drug's dose were increased or an alternative drug with a less favorable risk-benefit profile were chosen on the basis of what is actually an uninformative biomarker result, it may lead to worse treatment outcomes while contributing to increasing financial expenditures at a time when our health care system is least able to afford it.

As a result, we believe that testing recommendations supported by clinical validity alone adds confusion, not clarity, to the clinical decision making process, especially if the evidence is not clearly explicated or cited alongside the guidance. There may also be legal implications from the inclusion of testing recommendations in drug labels: could prescribers be liable for adverse events that may have been predicted by biomarker testing, regardless of the evidence base for the recommendations? If not, at what level of evidence should the "reasonableness" standard to test apply?

A multipronged approach should be used to address the current situation. At minimum, an explicit statement about the quality of clinical utility evidence for each testing recommendation should be presented in the labels, rather than charging clinicians with the task of extrapolating quality based on the data presented. Alongside this should be complete references to help patients and clinicians easily access the full studies for further assessment. More stringently, FDA could issue regulations to only include information about pharmacogenomic biomarkers if compelling clinical utility information has been generated, though exceptions should be made for drug-biomarker combinations associated with efficacy or safety endpoints of particular significance, such as clopidogrel and *CYP2C19*.

The perceived lack of incentive for biomarker test developers and pharmaceutical manufacturers to conduct robust studies describing the clinical utility of biomarker testing - which may demonstrate the lack of need for biomarker testing or reduce the number of patients eligible for a particular drug - poses another challenge. However, establishing the clinical utility of a biomarker for a particular high-risk patient population with high-quality studies may actually increase drug sales and biomarker test orders, as was the case with abacavir.[33] [34] To provide further incentive, the FDA could waive user fees and/or prioritize the review of a subsequent drug application for manufacturers who conduct robust pharmacogenomic trials in support of their approval applications. In addition,

governmental agencies can directly support trials investigating drug-biomarker combinations of particular clinical significance, as has been done for drugs for which manufacturers have lacked specific incentive to do so.[35]

In the meantime, physicians and other prescribers should be aware of the relative lack of evidence supporting many treatment recommendations contained within the labels and should instead scrutinize the primary literature supporting these recommendations before taking clinical action. Our finding that more than two-thirds of labels making testing recommendations do not contain convincing evidence supporting clinical utility of the biomarker further reinforces the need for skepticism.

Our study has several limitations. For each medication, we evaluated the earliest accessible label containing mention of each associated biomarker's polymorphisms in order to assess the evidence available to physicians at the time the biomarker was first included in the label. The initial labels of some of the drugs in our study were not accessible, which may have resulted in our evaluation of subsequent labels with updated pharmacogenomic evidence and recommendations. This may have resulted in an overestimate of the strength of the cited evidence. Moreover, studies that demonstrated a lack of clinical validity or utility of drug-biomarker relationships may not be included in the drug labels, which are drafted by manufacturers. However, since the FDA reviews and regulates the content of the labels, it should still require that major studies of such nature be included in subsequent labeling revisions. Additionally, we assessed the quality of clinical evidence only of information presented within the drug labels and additional full studies adequately referenced in these documents; we did not evaluate other sources of evidence, including additional primary literature or FDA medical reviews. As noted before, we decided to do this for two reasons: we believe that drug labels should be self-sufficient in presenting time-constrained prescribers with the evidence base and references supporting incorporation of biomarker testing into clinical practice, and the content presented in the labels is often used to inform other widely used tertiary-level drug information resources. It is worth noting that the results of our analysis are consistent with the three systematic reviews of drug-biomarker combinations that EGAPP has conducted and applied its evaluative framework.[23 24 25]

Also, while the FDA has generally required less robust evidence for approval of oncology drugs, with the goal of accelerating access for patients, we applied the same methodology in our evaluation of labels for biomarkers associated with these drugs to generate a standardized description of the evidence base available. Recent attention to the wide variation of clinical trial evidence submitted to the FDA to support the successful approval of novel agents[36], coupled with calls for more rigorous oncology trials to better reflect developments in treatment paradigms and clinical outcomes of different cancer types[37], further support the need to take a uniform approach to understanding the existing scientific landscape. Despite the decreased rigor required for approval of certain cancer medications, our analysis revealed that biomarkers for oncology drugs still demonstrated a significantly higher level of convincing evidence base in their labels than biomarkers for drugs treating other conditions.

CONCLUSION

There is reason to be enthusiastic about the potential of biomarkers to enhance clinical care and our analysis identified many examples of tests that have convincing evidence of their ability to meaningfully improve health care outcomes in patients with common conditions. However, other less evidence-based labeling recommendations highlight the need for clearer guidance on their optimal use. Until this problem is addressed, clinicians are left with the challenging task of navigating a sea of guidance with varying foundations of clinical support in pursuit of practicing clinically-sound and cost-conscious medicine, a challenge that will likely increase in cadence with the growth of the pharmacogenomics field.

FUTURE DIRECTIONS

Several steps can be taken going forward to expand upon our work and translate it to the improvement of health care delivery. In the near term, our study findings can be employed to inform clinical decision support systems to facilitate the optimal and judicious use of biomarker testing at the point of prescribing. Biomarker testing supported by convincing clinical utility evidence should be promoted and prioritized over the testing of biomarkers for which the drug-biomarker association is supported by incomplete clinical

utility evidence. In the latter case, the decision to use biomarker testing to inform pharmacotherapy should be made on an individual basis depending on the clinical context. We have thus far received correspondences from two prominent academic institutions in the United States who intend to carry our work forward in this very direction.

The FDA should also work in collaboration with drug manufacturers to ensure that the most robust evidence supporting each drug-biomarker association is described and adequately cited in the drug label, not only to efficiently inform physicians and other prescribers of a particular biomarker's clinical value, but also to help establish the landscape of overall evidence quality supporting each drug-biomarker combination, evidence described both within drug labels and elsewhere, in order to identify areas where research is most needed and incentives most likely to help close the knowledge gap surrounding the value of biomarker testing in guiding pharmacotherapy.

In addition, future discussions and analyses should explore whether the current EGAPP rating criteria should be customized for different types of diseases or different prevalence of adverse effects. For example, drug-biomarker combinations that are designed to address rare adverse events may be better evaluated with well-designed observational studies to generate evidence for clinical utility rather than relying exclusively on randomized controlled trials.

References

1 Timm R, Kaiser R, Lötsch J et al. Association of cyclophosphamide pharmacokinetics to polymorphic cytochrome P450 2C19. Pharmacogenomics J. 2005;5(6):365-373.

2 Kirchheiner J, Nickchen K, Bauer M et al. Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. Mol Psychiatry. 2004;9(5):442-473.

3 Lindh JD, Lundgren S, Holm L et al. Several-fold increase in risk of overanticoagulation by CYP2C9 mutations. Clin Pharmacol Ther. 2005;78(5):540-550.

4 Ingelman-Sundberg M, Sim SC, Gomez A et al. Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoepigenetic and clinical aspects. Pharmacol Ther. 2007;116(3):496-526.

5 Bristol-Myers Squibb Company. Coumadin. 2011. Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2011/009218s107lbl.pdf. Accessed May 14, 2014.

6 ViiV Healthcare. Ziagen. 2015. Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/020977s028,020978s032lbl.pdf. Accessed May 14, 2014.

7 Frueh FW, Amur S, Mummaneni P et al. Pharmacogenomic biomarker information in drug labels approved by the United States Food and Drug Administration: prevalence of related drug use. Pharmacotherapy. 2008;28(8):992-998.

8 US Food and Drug Administration.Table of pharmacogenomic biomarkers in drug labels.http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm. Accessed April 18, 2013.

9 Stanek EJ, Sanders CL and Frueh FW. Physician awareness and utilization of Food and Drug Administration (FDA)-approved labeling for pharmacogenomic testing information. J Pers Med 2013;3:111-123.

10 Davies KS. Physicians and their use of information: a survey comparison between the United States, Canada, and the United Kingdom. J Med Libr Assoc 2011;99(1):88-91.

11 Micromedex Mobile App General and Technical FAQs. Truven Health Analytics, 2014. Web. 13 June 2014. <http://micromedex.com/support/frequently-asked-questions/micromedex-20-mobile-app-iphone>.
12 Uhl K, Kennedy DL and Kweder SL. Risk management strategies in the Physicians' Desk Reference product labels for pregnancy category X drugs. Drug Safety 2002;25(12):885-892.
13 US Food and Drug Administration. An introduction to the improved FDA prescription drug labeling. Available from: http://www.fda.gov/downloads/Training/ForHealthProfessionals/UCM090796.pdf. Accessed May 30, 2015.
14 Stanek EJ, Sanders CL, Taber KA et al. Adoption of pharmacogenomic testing by US physicians: results of a nationwide survey. Clin Pharmacol Ther 2012;91(3):450-458.
15 Ong FS, Deignan JL, Kuo JZ et al. Clinical utility of pharmacogenetic biomarkers in cardiovascular therapeutics: a challenge for clinical implementation. Pharmacogenomics 2012; 13(4):465-475.
16 Kimmel SE, French B, Kasner SE et al. A pharmacogenetic versus a clinical algorithm for warfarin dosing. N Engl J Med 2013;369(24):2283-93.
17 Verhoef TI, Ragia G, de Boer A et al. A randomized trial of genotype-guided dosing of acenocoumarol and phenprocoumon. N Engl J Med 2013;369(24):2304-12.
18 Pirmohamed M, Burnside G, Eriksson N et al. A randomized trial of genotype-guided dosing of warfarin. NEJM 2013;369(24):2294-303.
19 Stergiopoulos K and Brown DL. Genotype-guided vs clinical dosing of warfarin and its analogues: meta-analysis of randomized clinical trials. JAMA Intern Med 2014; 10.1001/jamainternmed.2014.2368. [Epub ahead of print]
20 National Health Expenditure Projections 2012-2022. Centers for Medicare and Medicaid Services. Web. 21 June 2014. <http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/downloads/proj2012.pdf>.
21 The new science of personalized medicine: translating the promise into practice. PricewaterhouseCoopers 2009.

22 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 2009;11(1):3-14.
23 Matchar DB, Thakur ME, Grossman I et al. Testing for cytochrome P450 polymorphisms in adults with non-psychotic depression treated with selective serotonin reuptake inhibitors (SSRIs). AHRQ 2007. Publication No. 07-E002. <http://archive.ahrq.gov/clinic/tp/cyp450tp.htm>.
24 Palomaki et al. Can UGT1A1 genotyping reduce morbidity and mortality in patients with metastatic colorectal cancer treated with irinotecan? An evidence-based review. Genetics in Medicine 2009;11:21-34.
25 Lin JS, Webber EM, Senger CA. Systematic review of pharmacogenetic testing for predicting clinical benefit to anti-EGFR therapy in metastatic colorectal cancer. Am J Cancer Res 2011;1(5):650-662.
26 Leiter MP, Frank E, Matheson TJ. Demands, values, and burnout: relevance for physicians. Canadian Family Physician. 2009;55(12):1224-1225, 1225.e1-6.
27 Pimlott N. Who has time for family medicine? Canadian Family Physician. 2008;54(1):14-16.
28 Morrison I. The future of physician's time. Ann Intern Med. 2000;132(1):80-84.
29 US Food and Drug Administration. Drugs@FDA: FDA approved drug products. http://www.accessdata.fda.gov/scripts/cder/drugsatfda/. Accessed September 15, 2013.
30 PharmGKB: The Pharmacogenomics Knowledgebase. http://www.pharmgkb.org/. Accessed May 3, 2013.
31 DailyMed. http://dailymed.nlm.nih.gov/dailymed/about.cfm. Accessed December 3, 2013.
32 National Center for Biotechnology Information. U.S. National Library of Medicine, 2014. Web. 14 June 2014. <https://www.ncbi.nlm.nih.gov/pubmed/>.
33 Stratified medicine and pharmacogenomics. Priority Medicines for Europe and the World 2013 Update. World Health Organization, 2013. Web. 14 June 2014. <http://www.who.int/medicines/areas/priority_medicines/Ch7_4Stratified.pdf>.

34 Lai-Goldman M and Faruki H. Abacavir hypersensitivity: a model system for pharmacogenetic test adoption. Genet Med 2008;10(12):874-878.
35 Furberg CD, Wright Jr. JT, Davis BR et al. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: the antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). JAMA 2002;288(23):2981-2997.
36 Downing NS, Aminawung JA, Shah ND et al. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. JAMA 2014;311(4):368-377.
37 Hirsch BR, Califf RM, Cheng SK et al. Characteristics of oncology clinical trials: insights from a systematic analysis of ClinicalTrials.gov. JAMA Intern Med 2013;173(11):972-979.

Table 1 - Categorization of clinical evidence in drug labels

Quality	Clinical Validity[1]	Clinical Utility[2]
Convincing	High quality longitudinal cohort study Validated clinical decision rule Systematic review/meta-analysis of well-designed longitudinal cohort studies with homogeneity Systematic review/meta-analysis of randomized controlled trials or uncontrolled interventional trials (or sub-studies of these trials)‡ Randomized controlled trial or uncontrolled interventional trial‡	Systematic review/meta-analysis of randomized controlled trials showing consistency in results At least one large randomized controlled trial
Adequate	Systematic review of lower quality studies Case-control study with good reference standards Unvalidated Clinical Decision Rule Sub-study of large randomized controlled trial‡	Systematic review of randomized controlled trials with heterogeneity One or more controlled trials without randomization Systematic review of cohort studies with consistent results
Incomplete	Single case-control study using nonconsecutive cases or lacking in consistently applied reference standards Cohort or case-control study in which the reference standard is defined by the test or is not used systematically, or where study is not blinded Case report‡ or case series Unpublished and/or non-peer reviewed research, clinical laboratory, or manufacturer data Consensus guidelines No relevant studies or inability to classify study design based on description in drug label‡	Systematic review of non-randomized controlled trials, uncontrolled interventional trials‡, cohort studies, or case-control studies with heterogeneity Single cohort or case-control study Case series Unpublished and/or non-peer reviewed research, clinical laboratory, or manufacturer data One or more uncontrolled interventional trials‡ No relevant studies or inability to classify study design based on description in drug label‡

Table adapted from: 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 2009;11(1):3-14.

[1]Clinical validity of a genetic test is defined by the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) group as "its ability to accurately and reliably predict the clinically defined disorder or phenotype of interest."

[2]Clinical utility of a genetic test is defined by the EGAPP group as the "evidence of improved measurable clinical outcomes, and its usefulness and added value to patient management decision-making compared with current management without genetic testing."

[‡]These study designs were added to allow for classification of evidence that did not fit into the EGAPP categorization framework

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):1938-1944. Table 1. Categorization of Clinical Evidence in Drug Labels; p. 1940.

Table 2. Drug-biomarker combinations evaluated in study

	Drug	Biomarker		Drug	Biomarker
1	Abacavir	HLA-B*5701	61	Irinotecan	UGT1A1
2	Ado-Trastuzumab Emtansine	ERBB2 (HER2)	62	Isosorbide and Hydralazine	NAT1; NAT2
3	Aripiprazole	CYP2D6	63	Ivacaftor	CFTR (G551D)
4	Arsenic Trioxide	PML/RARα translocation	64	Lansoprazole	CYP2C19
5	Atomoxetine	CYP2D6	65	Lapatinib	Her2/neu
6	Atorvastatin	LDL receptor	66	Lenalidomide	Chromosome 5q
7	Azathioprine	TPMT	67	Letrozole	ER &/PgR receptor
8	Boceprevir	Interferon-lambda-3 (IL-28b)	68	Maraviroc	CCR5
9	Brentuximab Vedotin	CD30	69	Mercaptopurine	TPMT
10	Busulfan	Philadelphia chromosome	70	Metoprolol	CYP2D6
11	Capecitabine	DPD	71	Modafinil	CYP2D6
12	Carbamazepine	HLA-B*1502	72	Mycophenolic Acid	HGPRT
13	Carisoprodol	CYP2C19	73	Nefazodone	CYP2D6
14	Carvedilol	CYP2D6	74	Nilotinib (1)	Philadelphia chromosome
15	Celecoxib	CYP2C9	75	Nilotinib (2)	UGT1A1
16	Cetuximab (1)	EGFR	76	Nortriptyline	CYP2D6
17	Cetuximab (2)	KRAS	77	Omeprazole	CYP2C19
18	Cevimeline	CYP2D6	78	Panitumumab (1)	EGFR
19	Chlordiazepoxide and Amitriptyline	CYP2D6	79	Panitumumab (2)	KRAS
20	Chloroquine	G6PD	80	Pantoprazole	CYP2C19
21	Cisplatin	TPMT	81	Paroxetine	CYP2D6
22	Citalopram (1)	CYP2C19	82	Peginterferon alfa-2b	Interferon-lambda-3 (IL-28b)
23	Citalopram (2)	CYP2D6	83	Perphenazine	CYP2D6
24	Clobazam	CYP2C19	84	Pertuzumab	Her2/neu
25	Clomipramine	CYP2D6	85	Phenytoin	HLA-B*1502
26	Clopidogrel	CYP2C19	86	Pimozide	CYP2D6
27	Clozapine	CYP2D6	87	Prasugrel	CYP2C19
28	Codeine	CYP2D6	88	Pravastatin	ApoE2
29	Crizotinib	ALK	89	Propafenone	CYP2D6
30	Dapsone	G6PD	90	Propranolol	CYP2D6
31	Dasatinib	Philadelphia chromosome	91	Protriptyline	CYP2D6

32	Denileukin Diftitox	CD25	92	Quinidine	CYP2D6
33	Desipramine	CYP2D6	93	Rabeprazole	CYP2C19
34	Dexlansoprazole (1)	CYP2C19	94	Rasburicase	G6PD
35	Dexlansoprazole (2)	CYP1A2	95	Rifampin, Isoniazid, and Pyrazinamide	NAT1; NAT2
36	Dextromethorphan and Quinidine	CYP2D6	96	Risperidone	CYP2D6
37	Diazepam	CYP2C19	97	Sodium Phenylacetate and Sodium Benzoate	NAGS; CPS; ASS; OTC; ASL; ARG
38	Doxepin	CYP2D6	98	Sodium Phenylbutyrate	NAGS; CPS; ASS; OTC; ASL; ARG
39	Drospirenone and Ethinyl Estradiol	CYP2C19	99	Tamoxifen (1)	Estrogen receptor
40	Eltrombopag (1)	Factor V Leiden (FV)	100	Tamoxifen (2)	Factor V Leiden (FV)
41	Eltrombopag (2)	Antithrombin III deficiency (SERPINC1)	101	Tamoxifen (3)	Prothrombin mutations (F2)
42	Erlotinib	EGFR	102	Telaprevir	Interferon-lambda-3 (IL-28b)
43	Esomeprazole	CYP2C19	103	Terbinafine	CYP2D6
44	Everolimus	ERBB2 (HER2)	104	Tetrabenazine	CYP2D6
45	Exemestane	ER &/PgR receptor	105	Thioguanine	TPMT
46	Fluorouracil	DPD	106	Thioridazine	CYP2D6
47	Fluoxetine	CYP2D6	107	Ticagrelor	CYP2C19
48	Fluoxetine and Olanzapine	CYP2D6	108	Tolterodine	CYP2D6
49	Flurbiprofen	CYP2C9	109	Tositumomab	CD20 antigen
50	Fluvoxamine	CYP2D6	110	Tramadol and Acetaminophen	CYP2D6
51	Fulvestrant	Estrogen receptor	111	Trastuzumab	Her2/neu
52	Galantamine	CYP2D6	112	Tretinoin	PML/RARα translocation
53	Gefitinib	EGFR	113	Trimipramine	CYP2D6
54	Iloperidone	CYP2D6	114	Valproic Acid	NAGS; CPS; ASS; OTC; ASL; ARG
55	Imatinib (1)	C-Kit	115	Vemurafenib	BRAF
56	Imatinib (2)	Philadelphia chromosome	116	Venlafaxine	CYP2D6

57	Imatinib (3)	PDGFR (platelet-derived growth factor receptor) gene re-arrangements	117	Voriconazole	CYP2C19
58	Imatinib (4)	FIP1L1-PDGFRα fusion	118	Warfarin (1)	CYP2C9
59	Imipramine	CYP2D6	119	Warfarin (2)	VKORC1
60	Indacaterol	UGT1A1			

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):1938-1944. Appendix Table. Drug-Biomarker Combinations Evaluated in the Study.

Table 3. Features of evidence and recommendations in labels, stratified by therapeutic area

	No. (%)						
		Therapeutic Areas					
	All	Oncology (n=37)	Neuro-psychiatry (n=33)	GI (n=9)	ID (n=9)	CV (n=8)	Other (n=23)
Quality of Cited Evidence							
Clinical Validity							
Convincing	43 (36.1)	24 (64.9)	2 (6.1)	2 (22.2)	5 (55.6)	1 (12.5)	9 (39.1)
Adequate	6 (5.0)	4 (10.8)	1 (3.0)	0 (0)	0 (0)	1 (12.5)	0 (0)
Incomplete	70 (58.8)	9 (24.3)	30 (90.9)	7 (77.8)	4 (44.4)	6 (75)	14 (60.9)
Clinical Utility							
Convincing	18 (15.1)	14 (37.8)*	0 (0)‡	0 (0)	2 (22.2)	0 (0)	2 (8.7)
Adequate	0 (0)	0(0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Incomplete	101 (84.9)	23 (62.2)	33 (100)	9 (100)	7 (77.8)	8 (100)	21 (91.3)
Completeness of Study Citation							
Full	13 (10.9)	3 (8.1)	0 (0)	0 (0)	5 (55.6)	2 (25.0)	3 (13.0)
Partial	70 (58.8)	28 (75.7)	16 (48.5)	7 (77.8)	3 (33.3)	5 (62.5)	11 (47.8)
None	36 (30.3)	6 (16.2)	17 (51.5)	2 (22.2)	1 (11.1)	1 (12.5)	9 (39.1)
Presence of Testing Recommendation							
Based on mechanism of action	34 (28.6)	26 (70.3)	0 (0)	2 (22.2)	1 (11.1)	0 (0)	5 (21.7)
Based on drug-biomarker associations	27 (22.7)	3 (8.1)	14 (42.4)	1 (11.1)	1 (11.1)	1 (12.5)	7 (30.4)
Absent	58 (48.7)	8 (21.6)	19 (57.6)	6 (66.7)	7 (77.8)	7 (87.5)	11 (47.8)

GI = gastroenterology ID = infectious disease CV = cardiovascular
*p<0.001 compared to all other disease groupings (14/37 [38%] vs. 4/82 [5%])
‡p<0.001 compared to all other disease groupings (0/33 [0%] vs. 18/86 [21%])

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):1938-1944. Table 2. Features of Evidence and Recommendations in Labels Stratified by Therapeutic Area; p. 1941.

Table 4. Examples of evidence evaluation for clinical validity and utility

Drug Name	Biomarker	Excerpt from Drug Label of Studies Supporting Clinical Validity and/or Clinical Utility	Validity Study Design and Grade	Utility Study Design and Grade
Abacavir	*HLA-B*5701*	Validity and Utility: "CNA106030 (PREDICT-1), a randomized, double-blind study, evaluated the clinical utility of prospective HLA-B*5701 screening on the incidence of abacavir hypersensitivity reaction in abacavir-naive HIV-1-infected adults (n = 1,650). In this study, use of pre-therapy screening for the HLA-B*5701 allele and exclusion of subjects with this allele reduced the incidence of clinically suspected abacavir hypersensitivity reactions from 7.8% (66/847) to 3.4% (27/803)..."[a]	Randomized controlled trial	Randomized controlled trial
			Convincing	Convincing
Clopidogrel	*CYP2C19*	Validity: "The relationship between CYP2C19 genotype and Plavix treatment outcome was evaluated in retrospective analyses of Plavix-treated subjects in CHARISMA (n=4862) and TRITON-TIMI 38 (n=1477), and in several published cohort studies. In TRITON-TIMI 38 and the majority of the cohort studies, the combined group of patients with either intermediate or poor metabolizer status had a higher rate of cardiovascular events (death, myocardial infarction, and stroke) or stent thrombosis compared to extensive metabolizers. In CHARISMA and one cohort study, the increased event rate was observed only in poor metabolizers."[b]	Systematic review/meta-analysis of randomized controlled trials	No relevant studies
			Convincing	Incomplete

[a] Ziagen [package insert]. Research Triangle Park, NC: GlaxoSmithKline; 2008. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2008/020977s017,020978s020lbl.pdf
[b] Plavix [package insert]. Bridgewater, NJ: Bristol-Myers Squibb/Sanofi Pharmaceuticals Partnership; 2010. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2010/020839s042lbl.pdf

		Utility: No relevant data		
Carbamazepine	*HLA-B*1502*	Validity: "Retrospective case-control studies have found that in patients of Chinese ancestry there is a strong association between the risk of developing SJS/TEN with carbamazepine treatment and the presence of an inherited variant of the HLA-B gene, HLA-B*1502. The occurrence of higher rates of these reactions in countries with higher frequencies of this allele suggests that the risk may be increased in allele-positive individuals of any ethnicity."[c] Utility: No relevant data	Case-control study with good reference standards	No relevant studies
			Adequate	Incomplete
Fluorouracil	*DPD*	Validity: "One case of life-threatening systemic toxicity has been reported with the topical use of Efudex in a patient with DPD enzyme deficiency..."[d] Utility: No relevant data	Case report	No relevant studies
			Incomplete	Incomplete

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):1938-1944. Table 3. Examples of Evidence Evaluation for Clinical Validity and Utility; p.1942.

[c] Tegretol [package insert]. East Hanover, NJ: Novartis Pharmaceuticals Corporation; 2007. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2007/016608s098lbl.pdf
[d] Efudex [package insert]. Costa Mesa, CA: ICN Pharmaceuticals, Inc; 2004. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2004/16831slr047_efudex_lbl.pdf

Appendix Table 1. Individual drug-biomarker evaluations for clinical validity

Drug	Biomarker	Validity Study Design	Validity Grade
Abacavir	HLA-B*5701	Randomized controlled trial	convincing
Ado-Trastuzumab Emtansine	ERBB2 (HER2)	Randomized controlled trial	convincing
Aripiprazole	CYP2D6	Clinical laboratory data	incomplete
Arsenic Trioxide	PML/RARα translocation	Uncontrolled interventional trial	convincing
Atomoxetine	CYP2D6	Clinical laboratory data	incomplete
Atorvastatin	LDL receptor	High quality longitudinal cohort study‡	convincing
Azathioprine	TPMT	Uncontrolled interventional trial	convincing
Boceprevir	Interferon-lambda-3 (IL-28b)	Systematic review/meta-analysis of randomized controlled trials	convincing
Brentuximab Vedotin	CD30	No relevant studies	incomplete
Busulfan	Philadelphia chromosome	No relevant studies	incomplete
Capecitabine	DPD	No relevant studies	incomplete
Carbamazepine	HLA-B*1502	Case-control study with good reference standards	adequate
Carisoprodol	CYP2C19	No relevant studies	incomplete
Carvedilol	CYP2D6	Clinical laboratory data	incomplete
Celecoxib	CYP2C9	No relevant studies	incomplete
Cetuximab (1)	EGFR	Randomized controlled trial	convincing
Cetuximab (2)	KRAS	Systematic review/meta-analysis of randomized controlled trials	convincing
Cevimeline	CYP2D6	No relevant studies	incomplete
Chlordiazepoxide and Amitriptyline	CYP2D6	No relevant studies	incomplete
Chloroquine	G6PD	No relevant studies	incomplete
Cisplatin	TPMT	Case-control study with good reference standards	adequate
Citalopram (1)	CYP2C19	No relevant studies	incomplete
Citalopram (2)	CYP2D6	Clinical laboratory data	incomplete

Clobazam	CYP2C19	Clinical laboratory data	incomplete
Clomipramine	CYP2D6	No relevant studies	incomplete
Clopidogrel	CYP2C19	Systematic review/meta-analysis of randomized controlled trials	convincing
Clozapine	CYP2D6	No relevant studies	incomplete
Codeine	CYP2D6	No relevant studies	incomplete
Crizotinib	ALK	Systematic review/meta-analysis of uncontrolled interventional trials	convincing
Dapsone	G6PD	Inability to classify study design	incomplete
Dasatinib	Philadelphia chromosome	Systematic review/meta-analysis of uncontrolled interventional trials	convincing
Denileukin Diftitox	CD25	Randomized controlled trial	convincing
Desipramine	CYP2D6	No relevant studies	incomplete
Dexlansoprazole (1)	CYP2C19	Clinical laboratory data	incomplete
Dexlansoprazole (2)	CYP1A2	No relevant studies	incomplete
Dextromethorphan and Quinidine	CYP2D6	No relevant studies	incomplete
Diazepam	CYP2C19	No relevant studies	incomplete
Doxepin	CYP2D6	No relevant studies	incomplete
Drospirenone and Ethinyl Estradiol	CYP2C19	High quality longitudinal cohort study#	convincing
Eltrombopag (1)	Factor V Leiden (FV)	No relevant studies	incomplete
Eltrombopag (2)	Antithrombin III deficiency (SERPINC1)	No relevant studies	incomplete
Erlotinib	EGFR	Randomized controlled trial	convincing
Esomeprazole	CYP2C19	Clinical laboratory data	incomplete
Everolimus	ERBB2 (HER2)	Randomized controlled trial	convincing
Exemestane	ER &/PgR receptor	Randomized controlled trial	convincing
Fluorouracil	DPD	Case report	incomplete
Fluoxetine	CYP2D6	Clinical laboratory data	incomplete
Fluoxetine and Olanzapine	CYP2D6	Clinical laboratory data	incomplete
Flurbiprofen	CYP2C9	No relevant studies	incomplete

Fluvoxamine	CYP2D6	High quality longitudinal cohort study	convincing
Fulvestrant	Estrogen receptor	Systematic review/meta-analysis of randomized controlled trials	convincing
Galantamine	CYP2D6	High quality longitudinal cohort study	convincing
Gefitinib	EGFR	No relevant studies	incomplete
Iloperidone	CYP2D6	Clinical laboratory data	incomplete
Imatinib (1)	C-Kit	No relevant studies	incomplete
Imatinib (2)	Philadelphia chromosome	Systematic review/meta-analysis of uncontrolled interventional trials	convincing
Imatinib (3)	PDGFR (platelet-derived growth factor receptor) gene re-arrangements	Uncontrolled interventional trial	convincing
Imatinib (4)	FIP1L1-PDGFRα fusion	Uncontrolled interventional trial	convincing
Imipramine	CYP2D6	No relevant studies	incomplete
Indacaterol	UGT1A1	Clinical laboratory data	incomplete
Irinotecan	UGT1A1	Clinical laboratory data	incomplete
Isosorbide and Hydralazine	NAT1; NAT2	Clinical laboratory data	incomplete
Ivacaftor	CFTR (G551D)	Systematic review/meta-analysis of randomized controlled trials	convincing
Lansoprazole	CYP2C19	Clinical laboratory data#	incomplete
Lapatinib	Her2/neu	Randomized controlled trial	convincing
Lenalidomide	Chromosome 5q	Uncontrolled interventional trial	convincing
Letrozole	ER &/PgR receptor	Randomized controlled trial	convincing
Maraviroc	CCR5	Systematic review/meta-analysis of randomized controlled trials	convincing
Mercaptopurine	TPMT	Inability to classify study design	incomplete
Metoprolol	CYP2D6	Clinical laboratory data	incomplete
Modafinil	CYP2D6	Clinical laboratory data#	incomplete
Mycophenolic Acid	HGPRT	No relevant studies	incomplete

Nefazodone	CYP2D6	Clinical laboratory data#	incomplete
Nilotinib (1)	Philadelphia chromosome	Uncontrolled interventional trial	convincing
Nilotinib (2)	UGT1A1	Case-control study with good reference standards	adequate
Nortriptyline	CYP2D6	No relevant studies	incomplete
Omeprazole	CYP2C19	Systematic review/meta-analysis of randomized controlled trials*#	convincing
Panitumumab (1)	EGFR	Randomized controlled trial	convincing
Panitumumab (2)	KRAS	Systematic review/meta-analysis of randomized controlled trials	convincing
Pantoprazole	CYP2C19	Clinical laboratory data	incomplete
Paroxetine	CYP2D6	Clinical laboratory data#	incomplete
Peginterferon alfa-2b	Interferon-lambda-3 (IL-28b)	Systematic review/meta-analysis of randomized controlled trials	convincing
Perphenazine	CYP2D6	High quality longitudinal cohort study	incomplete
Pertuzumab	Her2/neu	Randomized controlled trial	convincing
Phenytoin	HLA-B*1502	Inability to classify study design	incomplete
Pimozide	CYP2D6	Clinical laboratory data	incomplete
Prasugrel	CYP2C19	Clinical laboratory data	incomplete
Pravastatin	ApoE2	Systematic review/meta-analysis of randomized controlled trials*	convincing
Propafenone	CYP2D6	Clinical laboratory data	incomplete
Propranolol	CYP2D6	Clinical laboratory data	incomplete
Protriptyline	CYP2D6	No relevant studies	incomplete
Quinidine	CYP2D6	No relevant studies	incomplete
Rabeprazole	CYP2C19	Clinical laboratory data	incomplete
Rasburicase	G6PD	Single case-control study lacking consistently applied reference standards	incomplete
Rifampin, Isoniazid, and Pyrazinamide	NAT1; NAT2	Inability to classify study design	incomplete
Risperidone	CYP2D6	Clinical laboratory data	incomplete

Sodium Phenylacetate and Sodium Benzoate	NAGS; CPS; ASS; OTC; ASL; ARG	High quality longitudinal cohort study	convincing
Sodium Phenylbutyrate	NAGS; CPS; ASS; OTC; ASL; ARG	Inability to classify study design	incomplete
Tamoxifen (1)	Estrogen receptor	Systematic review/meta-analysis of randomized controlled trials	convincing
Tamoxifen (2)	Factor V Leiden (FV)	Sub-study of large randomized controlled trial	adequate
Tamoxifen (3)	Prothrombin mutations (F2)	Sub-study of large randomized controlled trial	adequate
Telaprevir	Interferon-lambda-3 (IL-28b)	Systematic review/meta-analysis of randomized controlled trials	convincing
Terbinafine	CYP2D6	Clinical laboratory data$	incomplete
Tetrabenazine	CYP2D6	No relevant studies	incomplete
Thioguanine	TPMT	No relevant studies	incomplete
Thioridazine	CYP2D6	Clinical laboratory data	incomplete
Ticagrelor	CYP2C19	Sub-study of large randomized controlled trial	adequate
Tolterodine	CYP2D6	Clinical laboratory data	incomplete
Tositumomab	CD20 antigen	Uncontrolled interventional trial	convincing
Tramadol and Acetaminophen	CYP2D6	Clinical laboratory data	incomplete
Trastuzumab	Her2/neu	Randomized controlled trial	convincing
Tretinoin	PML/RARα translocation	Uncontrolled interventional trial	convincing
Trimipramine	CYP2D6	No relevant studies	incomplete
Valproic Acid	NAGS; CPS; ASS; OTC; ASL; ARG	No relevant studies	incomplete
Vemurafenib	BRAF	Randomized controlled trial	convincing
Venlafaxine	CYP2D6	Clinical laboratory data	incomplete
Voriconazole	CYP2C19	Clinical laboratory data	incomplete
Warfarin (1)	CYP2C9	Systematic review/meta-analysis of randomized controlled trials	convincing
Warfarin (2)	VKORC1	High quality longitudinal	convincing

		cohort study	

Appendix Table 2. Individual drug-biomarker evaluations for clinical utility

Drug	Biomarker	Utility Study Design	Utility Grade
Abacavir	HLA-B*5701	Randomized controlled trial	convincing
Ado-Trastuzumab Emtansine	ERBB2 (HER2)	Randomized controlled trial	convincing
Aripiprazole	CYP2D6	No relevant studies	incomplete
Arsenic Trioxide	PML/RARα translocation	Uncontrolled interventional trial	incomplete
Atomoxetine	CYP2D6	No relevant studies	incomplete
Atorvastatin	LDL receptor	Uncontrolled interventional trial‡	incomplete
Azathioprine	TPMT	No relevant studies	incomplete
Boceprevir	Interferon-lambda-3 (IL-28b)	No relevant studies	incomplete
Brentuximab Vedotin	CD30	No relevant studies	incomplete
Busulfan	Philadelphia chromosome	No relevant studies	incomplete
Capecitabine	DPD	No relevant studies	incomplete
Carbamazepine	HLA-B*1502	No relevant studies	incomplete
Carisoprodol	CYP2C19	No relevant studies	incomplete
Carvedilol	CYP2D6	No relevant studies	incomplete
Celecoxib	CYP2C9	No relevant studies	incomplete
Cetuximab (1)	EGFR	Randomized controlled trial	convincing
Cetuximab (2)	KRAS	No relevant studies	incomplete
Cevimeline	CYP2D6	No relevant studies	incomplete
Chlordiazepoxide and Amitriptyline	CYP2D6	No relevant studies	incomplete
Chloroquine	G6PD	No relevant studies	incomplete
Cisplatin	TPMT	No relevant studies	incomplete
Citalopram (1)	CYP2C19	No relevant studies	incomplete
Citalopram (2)	CYP2D6	No relevant studies	incomplete
Clobazam	CYP2C19	No relevant studies	incomplete
Clomipramine	CYP2D6	No relevant studies	incomplete

Clopidogrel	CYP2C19	No relevant studies	incomplete
Clozapine	CYP2D6	No relevant studies	incomplete
Codeine	CYP2D6	No relevant studies	incomplete
Crizotinib	ALK	Systematic review of uncontrolled interventional trials with heterogeneity	incomplete
Dapsone	G6PD	No relevant studies	incomplete
Dasatinib	Philadelphia chromosome	Systematic review of uncontrolled interventional trials with heterogeneity	incomplete
Denileukin Diftitox	CD25	Randomized controlled trial	convincing
Desipramine	CYP2D6	No relevant studies	incomplete
Dexlansoprazole (1)	CYP2C19	No relevant studies	incomplete
Dexlansoprazole (2)	CYP1A2	No relevant studies	incomplete
Dextromethorphan and Quinidine	CYP2D6	No relevant studies	incomplete
Diazepam	CYP2C19	No relevant studies	incomplete
Doxepin	CYP2D6	No relevant studies	incomplete
Drospirenone and Ethinyl Estradiol	CYP2C19	No relevant studies	incomplete
Eltrombopag (1)	Factor V Leiden (FV)	No relevant studies	incomplete
Eltrombopag (2)	Antithrombin III deficiency (SERPINC1)	No relevant studies	incomplete
Erlotinib	EGFR	Randomized controlled trial	convincing
Esomeprazole	CYP2C19	No relevant studies	incomplete
Everolimus	ERBB2 (HER2)	Randomized controlled trial	convincing
Exemestane	ER &/PgR receptor	Randomized controlled trial	convincing
Fluorouracil	DPD	No relevant studies	incomplete
Fluoxetine	CYP2D6	No relevant studies	incomplete
Fluoxetine and Olanzapine	CYP2D6	No relevant studies	incomplete
Flurbiprofen	CYP2C9	No relevant studies	incomplete
Fluvoxamine	CYP2D6	No relevant studies	incomplete

Fulvestrant	Estrogen receptor	Systematic review/meta-analysis of randomized controlled trials showing consistency in results	convincing
Galantamine	CYP2D6	No relevant studies	incomplete
Gefitinib	EGFR	No relevant studies	incomplete
Iloperidone	CYP2D6	No relevant studies	incomplete
Imatinib (1)	C-Kit	No relevant studies	incomplete
Imatinib (2)	Philadelphia chromosome	Systematic review of uncontrolled interventional trials with heterogeneity	incomplete
Imatinib (3)	PDGFR (platelet-derived growth factor receptor) gene re-arrangements	Uncontrolled interventional trial	incomplete
Imatinib (4)	FIP1L1-PDGFRα fusion	Uncontrolled interventional trial	incomplete
Imipramine	CYP2D6	No relevant studies	incomplete
Indacaterol	UGT1A1	No relevant studies	incomplete
Irinotecan	UGT1A1	No relevant studies	incomplete
Isosorbide and Hydralazine	NAT1; NAT2	No relevant studies	incomplete
Ivacaftor	CFTR (G551D)	Systematic review/meta-analysis of randomized controlled trials showing consistency in results	convincing
Lansoprazole	CYP2C19	No relevant studies	incomplete
Lapatinib	Her2/neu	Randomized controlled trial	convincing
Lenalidomide	Chromosome 5q	Uncontrolled interventional trial	incomplete
Letrozole	ER &/PgR receptor	Randomized controlled trial	convincing

Maraviroc	CCR5	Systematic review/meta-analysis of randomized controlled trials showing consistency in results	convincing
Mercaptopurine	TPMT	No relevant studies	incomplete
Metoprolol	CYP2D6	No relevant studies	incomplete
Modafinil	CYP2D6	No relevant studies	incomplete
Mycophenolic Acid	HGPRT	No relevant studies	incomplete
Nefazodone	CYP2D6	No relevant studies	incomplete
Nilotinib (1)	Philadelphia chromosome	Uncontrolled interventional trial	incomplete
Nilotinib (2)	UGT1A1	No relevant studies	incomplete
Nortriptyline	CYP2D6	No relevant studies	incomplete
Omeprazole	CYP2C19	No relevant studies	incomplete
Panitumumab (1)	EGFR	Randomized controlled trial	convincing
Panitumumab (2)	KRAS	No relevant studies	incomplete
Pantoprazole	CYP2C19	No relevant studies	incomplete
Paroxetine	CYP2D6	No relevant studies	incomplete
Peginterferon alfa-2b	Interferon-lambda-3 (IL-28b)	No relevant studies	incomplete
Perphenazine	CYP2D6	No relevant studies	incomplete
Pertuzumab	Her2/neu	Randomized controlled trial	convincing
Phenytoin	HLA-B*1502	No relevant studies	incomplete
Pimozide	CYP2D6	No relevant studies	incomplete
Prasugrel	CYP2C19	No relevant studies	incomplete
Pravastatin	ApoE2	Systematic review/meta-analysis of randomized controlled trials showing consistency in results*	convincing
Propafenone	CYP2D6	No relevant studies	incomplete
Propranolol	CYP2D6	No relevant studies	incomplete
Protriptyline	CYP2D6	No relevant studies	incomplete

Quinidine	CYP2D6	No relevant studies	incomplete
Rabeprazole	CYP2C19	No relevant studies	incomplete
Rasburicase	G6PD	No relevant studies	incomplete
Rifampin, Isoniazid, and Pyrazinamide	NAT1; NAT2	No relevant studies	incomplete
Risperidone	CYP2D6	No relevant studies	incomplete
Sodium Phenylacetate and Sodium Benzoate	NAGS; CPS; ASS; OTC; ASL; ARG	Single cohort study	incomplete
Sodium Phenylbutyrate	NAGS; CPS; ASS; OTC; ASL; ARG	No relevant studies	incomplete
Tamoxifen (1)	Estrogen receptor	Systematic review/meta-analysis of randomized controlled trials showing consistency in results	convincing
Tamoxifen (2)	Factor V Leiden (FV)	No relevant studies	incomplete
Tamoxifen (3)	Prothrombin mutations (F2)	No relevant studies	incomplete
Telaprevir	Interferon-lambda-3 (IL-28b)	No relevant studies	incomplete
Terbinafine	CYP2D6	No relevant studies	incomplete
Tetrabenazine	CYP2D6	No relevant studies	incomplete
Thioguanine	TPMT	No relevant studies	incomplete
Thioridazine	CYP2D6	No relevant studies	incomplete
Ticagrelor	CYP2C19	No relevant studies	incomplete
Tolterodine	CYP2D6	No relevant studies	incomplete
Tositumomab	CD20 antigen	Uncontrolled interventional trial	incomplete
Tramadol and Acetaminophen	CYP2D6	No relevant studies	incomplete
Trastuzumab	Her2/neu	Randomized controlled trial	convincing
Tretinoin	PML/RARα translocation	Uncontrolled interventional trial	incomplete
Trimipramine	CYP2D6	No relevant studies	incomplete

Valproic Acid	NAGS; CPS; ASS; OTC; ASL; ARG	No relevant studies	incomplete
Vemurafenib	BRAF	Randomized controlled trial	convincing
Venlafaxine	CYP2D6	No relevant studies	incomplete
Voriconazole	CYP2C19	No relevant studies	incomplete
Warfarin (1)	CYP2C9	No relevant studies	incomplete
Warfarin (2)	VKORC1	No relevant studies	incomplete

Appendix Table 3. Individual drug-biomarker evaluations for completeness of citation of supporting studies and presence of treatment recommendations

Drug	Biomarker	Completeness of Description of Supporting Evidence	Presence of Testing Recommendation
Abacavir	HLA-B*5701	Full	Present
Ado-Trastuzumab Emtansine	ERBB2 (HER2)	partial	Present
Aripiprazole	CYP2D6	partial	absent
Arsenic Trioxide	PML/RARα translocation	partial	Present
Atomoxetine	CYP2D6	partial	Present
Atorvastatin	LDL receptor	partial	Present
Azathioprine	TPMT	Full	Present
Boceprevir	Interferon-lambda-3 (IL-28b)	Full	absent
Brentuximab Vedotin	CD30	none	Present
Busulfan	Philadelphia chromosome	partial	Present
Capecitabine	DPD	none	Present
Carbamazepine	HLA-B*1502	partial	Present
Carisoprodol	CYP2C19	none	absent
Carvedilol	CYP2D6	partial	absent
Celecoxib	CYP2C9	none	absent
Cetuximab (1)	EGFR	partial	Present
Cetuximab (2)	KRAS	partial	Present
Cevimeline	CYP2D6	none	absent
Chlordiazepoxide and Amitriptyline	CYP2D6	none	absent
Chloroquine	G6PD	none	absent
Cisplatin	TPMT	partial	absent
Citalopram (1)	CYP2C19	none	Present
Citalopram (2)	CYP2D6	none	absent
Clobazam	CYP2C19	partial	Present
Clomipramine	CYP2D6	none	absent
Clopidogrel	CYP2C19	Full	Present

Clozapine	CYP2D6	none	absent
Codeine	CYP2D6	none	Present
Crizotinib	ALK	partial	Present
Dapsone	G6PD	partial	Present
Dasatinib	Philadelphia chromosome	partial	Present
Denileukin Diftitox	CD25	partial	Present
Desipramine	CYP2D6	none	absent
Dexlansoprazole (1)	CYP2C19	partial	absent
Dexlansoprazole (2)	CYP1A2	none	absent
Dextromethorphan and Quinidine	CYP2D6	none	Present
Diazepam	CYP2C19	none	absent
Doxepin	CYP2D6	none	absent
Drospirenone and Ethinyl Estradiol	CYP2C19	partial	absent
Eltrombopag (1)	Factor V Leiden (FV)	none	absent
Eltrombopag (2)	Antithrombin III deficiency (SERPINC1)	none	absent
Erlotinib	EGFR	partial	Present
Esomeprazole	CYP2C19	partial	absent
Everolimus	ERBB2 (HER2)	partial	Present
Exemestane	ER &/PgR receptor	partial	Present
Fluorouracil	DPD	partial	Present
Fluoxetine	CYP2D6	partial	absent
Fluoxetine and Olanzapine	CYP2D6	partial	absent
Flurbiprofen	CYP2C9	none	absent
Fluvoxamine	CYP2D6	partial	absent
Fulvestrant	Estrogen receptor	partial	Present
Galantamine	CYP2D6	partial	absent
Gefitinib	EGFR	none	absent
Iloperidone	CYP2D6	partial	Present
Imatinib (1)	C-Kit	none	absent

Imatinib (2)	Philadelphia chromosome	partial	Present
Imatinib (3)	PDGFR (platelet-derived growth factor receptor) gene re-arrangements	partial	Present
Imatinib (4)	FIP1L1-PDGFRα fusion	partial	Present
Imipramine	CYP2D6	none	absent
Indacaterol	UGT1A1	partial	absent
Irinotecan	UGT1A1	partial	absent
Isosorbide and Hydralazine	NAT1; NAT2	none	absent
Ivacaftor	CFTR (G551D)	partial	Present
Lansoprazole	CYP2C19	none	absent
Lapatinib	Her2/neu	partial	Present
Lenalidomide	Chromosome 5q	partial	Present
Letrozole	ER &/PgR receptor	partial	Present
Maraviroc	CCR5	Full	Present
Mercaptopurine	TPMT	none	Present
Metoprolol	CYP2D6	partial	absent
Modafinil	CYP2D6	none	Present
Mycophenolic Acid	HGPRT	none	Present
Nefazodone	CYP2D6	partial	Present
Nilotinib (1)	Philadelphia chromosome	partial	Present
Nilotinib (2)	UGT1A1	partial	absent
Nortriptyline	CYP2D6	none	absent
Omeprazole	CYP2C19	partial	Present
Panitumumab (1)	EGFR	partial	Present
Panitumumab (2)	KRAS	partial	Present
Pantoprazole	CYP2C19	partial	absent
Paroxetine	CYP2D6	partial	Present

Peginterferon alfa-2b	Interferon-lambda-3 (IL-28b)	Full	absent
Perphenazine	CYP2D6	partial	absent
Pertuzumab	Her2/neu	partial	Present
Phenytoin	HLA-B*1502	none	Present
Pimozide	CYP2D6	partial	Present
Prasugrel	CYP2C19	partial	absent
Pravastatin	ApoE2	partial	Present
Propafenone	CYP2D6	partial	absent
Propranolol	CYP2D6	partial	absent
Protriptyline	CYP2D6	none	absent
Quinidine	CYP2D6	none	absent
Rabeprazole	CYP2C19	partial	absent
Rasburicase	G6PD	partial	Present
Rifampin, Isoniazid, and Pyrazinamide	NAT1; NAT2	partial	absent
Risperidone	CYP2D6	partial	absent
Sodium Phenylacetate and Sodium Benzoate	NAGS; CPS; ASS; OTC; ASL; ARG	partial	Present
Sodium Phenylbutyrate	NAGS; CPS; ASS; OTC; ASL; ARG	partial	Present
Tamoxifen (1)	Estrogen receptor	Full	Present
Tamoxifen (2)	Factor V Leiden (FV)	Full	absent
Tamoxifen (3)	Prothrombin mutations (F2)	Full	absent
Telaprevir	Interferon-lambda-3 (IL-28b)	Full	absent
Terbinafine	CYP2D6	partial	absent
Tetrabenazine	CYP2D6	none	Present
Thioguanine	TPMT	none	absent
Thioridazine	CYP2D6	partial	Present
Ticagrelor	CYP2C19	Full	absent

Tolterodine	CYP2D6	partial	absent
Tositumomab	CD20 antigen	partial	Present
Tramadol and Acetaminophen	CYP2D6	partial	absent
Trastuzumab	Her2/neu	partial	Present
Tretinoin	PML/RARα translocation	partial	Present
Trimipramine	CYP2D6	none	absent
Valproic Acid	NAGS; CPS; ASS; OTC; ASL; ARG	none	Present
Vemurafenib	BRAF	partial	Present
Venlafaxine	CYP2D6	partial	absent
Voriconazole	CYP2C19	partial	absent
Warfarin (1)	CYP2C9	Full	Present
Warfarin (2)	VKORC1	Full	Present