

A formal risk-benefit framework for genomic tests: Facilitating the appropriate translation of genomics into clinical practice

David L. Veenstra, PharmD, PhD^{1,4}, Joshua A. Roth, MHA¹, Louis P. Garrison, Jr, PhD¹,
Scott D. Ramsey, MD, PhD^{3,4}, and Wylie Burke, MD, PhD^{2,4}

Purpose: Evaluation of genomic tests is often challenging because of the lack of direct evidence of clinical benefit compared with usual care and unclear evidence requirements. To address these issues, this study presents a risk-benefit framework for assessing the health-related utility of genomic tests. **Methods:** We incorporated approaches from a variety of established fields including decision science, outcomes research, and health technology assessment to develop the framework. Additionally, we considered genomic test stakeholder perspectives and case studies. **Results:** We developed a three-tiered framework: first, we use decision-analytic modeling techniques to synthesize data, project incidence of clinical events, and assess uncertainty. Second, we defined the health-related utility of genomic tests as improvement in health outcomes as measured by clinical event rates, life expectancy, and quality-adjusted life-years. Finally, we displayed results using a risk-benefit policy matrix to facilitate the interpretation and implementation of findings from these analyses. **Conclusion:** A formal risk-benefit framework may accelerate the utilization and practice-based evidence development of genomic tests that pose low risk and offer plausible clinical benefit, while discouraging premature use of tests that provide little benefit or pose significant health risks compared with usual care. *Genet Med* 2010;12(11):686–693.

Key Words: risk-benefit, genomic, translation, stakeholder, model

Few genomic testing technologies have reached routine clinical practice or been incorporated into clinical guidelines to date.^{1–4} Nonetheless, a multitude of genomic tests are marketed to consumers and physicians, and genome-wide assays are available to consumers for several hundred dollars.⁵ These assays, coupled with the rapid growth of somatic gene expression profiling in oncology, present a significant challenge to clinicians and policy makers seeking to establish clinical practices that maximize benefit for patients while minimizing harm.

The efficient and appropriate translation of genomic discoveries into clinical practice is particularly challenging because of an interrelated combination of factors.⁶ First, there is a notable

lack of comparative effectiveness data for genomic applications because of regulatory and reimbursement policies that neither require nor incentivize investment in such studies.^{7–9} Consequently, although randomized trials have been initiated for select genomic applications such as *CYP2C9/VKORC1* testing with warfarin therapy,^{10,11} *CYP2D6* testing with antidepressant use,¹² and gene expression profiling in breast cancer treatment,¹³ there are generally few prospective comparative genomic tests evaluations planned or underway.¹⁴

Second, the ease of market access for genomic tests makes the aforementioned lack of evidence more problematic.^{15,16} For example, when investigators from the National Institute of Mental Health reported an association between two genetic variants and suicidal ideation in patients taking citalopram,¹⁷ within a week, a genomic testing company announced plans to offer testing to “help to reduce a recently announced spike in suicide rates among US youth.”¹⁸ This situation is partly related to regulatory policy, but is also related to the fact that providing information about genomic susceptibilities does not require specialized medical facilities or training, and involves very little direct risk of immediate harm to patients.

Finally, there is a lack of consensus on evidence requirements or thresholds for genomic test evaluation.¹⁹ Some stakeholders accept the findings of retrospective analyses and clinical plausibility, whereas others expect controlled clinical trial data.^{20,21} For example, in the case of the anticoagulant warfarin, variants of the genes *CYP2C9* and *VKORC1* are clearly associated with lower dose requirements, but no study to date has definitively demonstrated that using this information improves patient outcomes.²² Alternatively, warfarin patients concomitantly taking amiodarone also require lower warfarin dosing (because of inhibition of *CYP2C9*), and doing so is considered standard of care.²³ This lack of consistency in evidence requirements, in addition to the other factors outlined above, creates a roadblock on the translational pathway for genomic tests.

The Secretary’s Advisory Committee on Genetics, Health, and Society recently issued a report¹⁵ emphasizing the importance of assessing and weighing potential harm against potential benefit, so that patients do not inadvertently forgo real benefit because of small or hypothetical harms. Additionally, regulatory authorities have shown heightened interest in the use of quantitative approaches to assess risk-benefit tradeoffs for pharmaceuticals.^{24–31} A recent Institute of Medicine study advised that Food and Drug Administration (FDA) “develop and continually improve a systematic approach to risk-benefit analysis.”³² FDA is currently evaluating various approaches to incorporate risk-benefit analyses into their assessment processes. Although approaches have been developed to incorporate indirect evidence (e.g., noncomparative data) in a semiquantitative fashion, and decision-analytic techniques are beginning to be applied in the assessment of genomic tests,^{4,33} quantitative assessment of risk-benefit tradeoffs, and the uncertainty sur-

From the ¹Department of Pharmacy; ²Department of Bioethics and Humanities, University of Washington; ³Division of Population Health Sciences, Fred Hutchinson Cancer Research Center; and ⁴Institute for Public Health Genomics, University of Washington, Seattle, Washington.

David L. Veenstra, PharmD, PhD, Department of Pharmacy, University of Washington, Box 357630, Seattle, WA 98195. E-mail: veenstra@u.washington.edu.

Disclaimer: The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention.

Disclosure: The authors declare no conflict of interest.

Submitted for publication April 24, 2010.

Accepted for publication June 24, 2010.

Published online ahead of print August 30, 2010.

DOI: 10.1097/GIM.0b013e3181eff533

rounding them, have not been explicitly included in genomic test evidence recommendations to date.

We believe that there is a significant opportunity to use existing decision modeling methods to synthesize genomic, clinical, epidemiological, and patient outcome data to explicitly evaluate risk and benefit trade-offs of genomic tests, and the uncertainty surrounding their utility. The objective of this study was to develop a systematic and comprehensive approach to help clinicians and policy-makers estimate health outcomes of genomic testing in the absence of definitive data. The novel aspect of the risk-benefit framework described in this article is the synthesis of approaches from a variety of fields to systematically and quantitatively evaluate the risk-benefit profile of genomic tests—the use of decision modeling, the projection of multiple clinical outcomes (including quality-adjusted life-years [QALYs] as a summary measure of clinical utility), and a recommendation framework that enables utilization of the information generated. These estimates are intended to help guide decisions about clinical test use and coverage and provide a framework for encouraging practice-based evidence development for tests with plausible net health benefit.

METHODS

The risk-benefit framework presented herein is based on work from the fields of decision science, outcomes research, and health technology assessment. Traditional evidence-based processes have generally relied on direct evidence of clinical utility (e.g., data from randomized controlled trials). Recently, however, advisory bodies have recognized that direct evidence will not always be available to answer questions of interest. For example, the U.S. Preventive Services Task Force (USPSTF) developed an approach for evaluating indirect evidence with a focus on evaluating net health benefit, and the uncertainty around estimates.^{34,35} The Task Force constructs a “chain of evidence” within an analytic framework and assesses the level of certainty based on specific questions. If the certainty of net benefit is moderate or high, the magnitude of benefit is assessed, and modeled event rates are provided in an outcomes table. For example, Nelson et al.³⁶ used this approach to evaluate BRCA mutation testing for breast and ovarian cancer susceptibility, although a summary measure of the net health benefit was not determined.

More recently, the U.S. Centers for Disease Control and Prevention has sponsored the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) initiative.³⁷ EGAPP’s methods are analogous to that of the USPSTF and involve use of an analytic framework to assess indirect evidence. Three of the seven EGAPP evidence reports commissioned to date have explicitly conducted decision-analytic modeling.³⁸ In two cases, evidence supporting a valid association between variants and clinical outcomes was lacking, and the models were used in an exploratory capacity.^{39,40} In the other case, the model was used to assess efficiency of case detection but not patient outcomes.⁴¹ A summary measure of net health benefit was not calculated in any of these cases.

The private sector has also pursued analogous, evidence-based approaches. Notably, the BlueCross BlueShield Association’s Technology Evaluation Center (TEC) has conducted extensive evidence-based evaluations of genomic tests.⁴² The TEC uses five criteria to evaluate health technologies such as genomic tests: (1) it must have regulatory approval, (2) the evidence must permit conclusions regarding its effect on health outcomes, (3) it must result in an improvement in net health outcomes, (4) it must be at least as good as current alternatives,

and (5) its benefits must be attainable outside of the investigational setting. Quantitative evaluation of indirect evidence has not been used for TEC assessments to date.

In summary, although approaches to date have incorporated various aspects of a quantitative risk-benefit framework, they have not included a formal and explicit approach to assessing indirect evidence, a summary measure of risk-benefit, and a decision-making framework that synthesizes this information. Below, we propose a quantitative risk-benefit approach that incorporates these aspects within a single framework. We used stakeholder feedback and previous experience with case studies and regulatory science to inform development of the framework.^{19,43–45}

Decision-analytic framework

Decision-analytic modeling provides an explicit framework for evaluating technologies by incorporating data from various sources in a quantitative and transparent fashion and comparing the likely results of technology use versus the next best alternative. By assessing the incremental outcomes compared with the next best alternative (e.g., no genomic testing), the “opportunity cost” of genomic testing can be captured. Weinstein and Fineberg⁴⁶ characterize the decision-analytic approach as (1) identifying and bounding the decision problem, (2) structuring the decision problem over time, (3) characterizing the information needed to inform the structure, and (4) choosing a preferred course of action. This approach is advantageous in that there is an explicit framework for evaluating risks and benefits, decision makers must identify quantitative estimates of risks and benefits, the approach can be applied to a wide variety of technologies, and complexity and timing of analyses can be suited to the decision-making task.²⁴

To illustrate the decision-modeling process, we consider a hypothetical cohort of patients initiating long-term warfarin therapy for the prevention of thromboembolic events. During the warfarin initiation period, determination of the dose required to achieve an optimal level of anticoagulation can be challenging. Clinicians monitor the international normalized ratio (INR), a measure of anticoagulation status that can serve as a surrogate marker for adverse events. INR values between 2 and 3 are considered within therapeutic range for most patients—INR values above 3 are associated with higher risk of serious bleeding events, whereas INR values below 2 are associated with increased risk of thromboembolic events. Most patients are initiated on 5 mg warfarin per day, and clinical and demographic variables that indicate warfarin sensitivity such as older age, drug interactions, or comorbidities are used to adjust doses downward. Information about the patients’ *CYP2C9* and *VKORC1* gene status (hereafter referred to as “genotype-guided” dosing) also could be incorporated in the initial dose selection. Below, we demonstrate how decision modeling can be used to quantitatively evaluate the risks and benefits of each approach based on an analysis conducted as part of this risk-benefit framework project, as well as the results of a previously published warfarin decision analytic model.^{43,47}

Decision structure, data sources, and outcomes

At the core of the risk-benefit framework is what could be described as a clinical disease-based model. The goal of this approach is to incorporate relevant clinical effects attributable to a genomic test and subsequent actions to estimate impact on patient outcomes. A schematic of the process is depicted in Figure 1.

Consider this approach in the context of genomic testing to guide warfarin therapy described earlier. First, clinicians re-

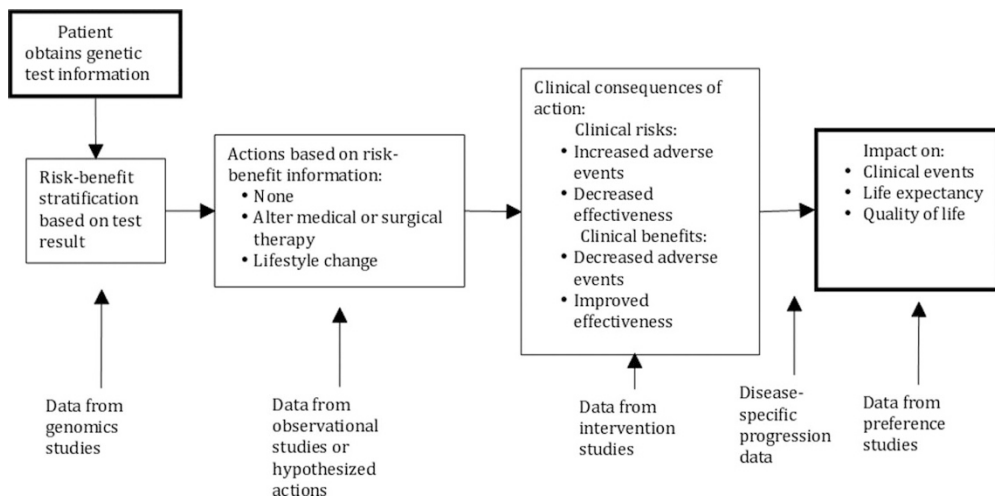


Fig. 1. Schematic diagram of disease-based model.

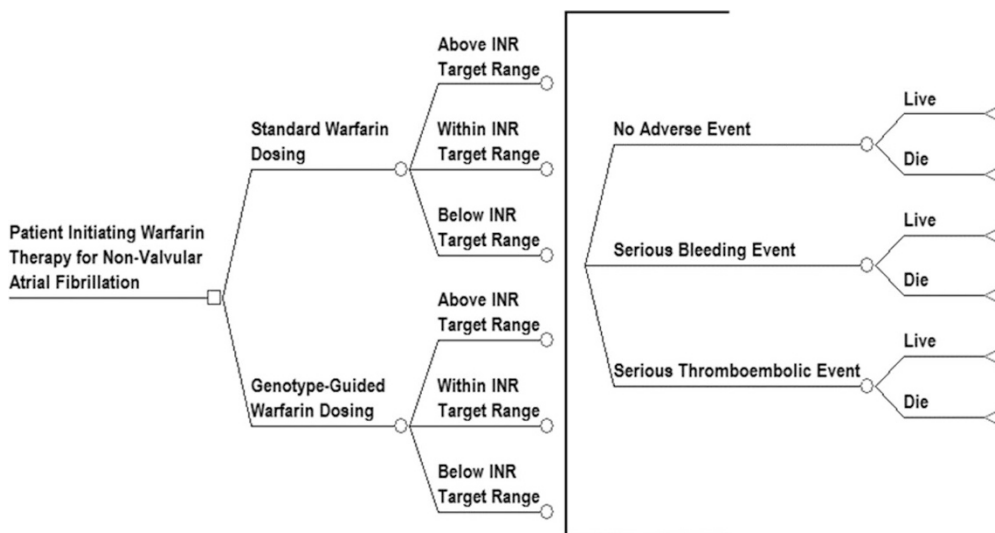


Fig. 2. Warfarin pharmacogenomics decision tree.

ceive genomic test results reporting patients' *CYP2C9* and *VKORC1* genomic status. Next, informed by the test results, an estimated initial dose is calculated and warfarin therapy is administered—assuming the clinician and patient agree with the suggested dosing. During the subsequent weeks, clinicians will monitor INR and adjust warfarin dose in response. The goal is to achieve stable INR values in therapeutic range, and a standard measure of anticoagulation management success is the time in therapeutic range over the first month (or months) of warfarin therapy.⁴⁸

We developed a risk-benefit analysis for warfarin pharmacogenomic testing based on extensive interaction with various stakeholders, particularly practicing anticoagulation clinicians.⁴⁷ Because of their familiarity with INR as an outcome, clinicians indicated that a model that projected bleed and clot events based on INR during the first 1–3 months of warfarin therapy would be most useful to assess the potential net benefit of testing (Fig. 2). The probabilities of achieving different levels of INR control ideally would be informed by the results of comparative, randomized clinical

trials. In this instance, we used results from the highest quality randomized controlled trial available to date conducted by Anderson et al. ($N = 200$).⁴⁹ Additionally, the relationship between time in INR range and the risk of clinical events can be derived from longitudinal cohort studies as was done with data from van Walraven et al.⁵⁰ in the warfarin model under consideration. These probabilities are then multiplied to compare the overall likelihood of having specific events within each dosing strategy.

We estimated in a cohort of 10,000 patients observed for the first month of warfarin therapy, approximately 44 and 45 patients would experience serious bleeding events and approximately 27 and 28 would experience serious thromboembolic events in the genotype-guided and standard dosing strategies, respectively.⁴⁷

Uncertainty: Scenario and sensitivity analysis

The decision structure described above does not allow traditional statistical analyses and hypothesis testing because not all

data are derived from the same study, nor typically obtained at the patient level. However, it is possible to explicitly evaluate uncertainty—particularly related to lack of data. To accomplish this, scenario analyses can be conducted in which model inputs are varied over plausible ranges, and the impact on results assessed—for instance in “most likely,” “best case,” and “worst case” scenarios. Each model input can be varied individually to identify inputs that drive the analysis and are associated with the greatest uncertainty in the results.

For example, in the decision tree depicted in Figure 2, uncertainty about the proportion of time patients spend within the target INR range in the first month of warfarin therapy could be explored by examining the modeled outcomes of a plausible range of values. Perhaps in the “most likely,” “best,” and “worst” scenarios, patients are within the INR target range for 66%, 82%, and 50% of the time, respectively. Downstream outcomes of these times in target INR range can be modeled to see how use of genotype-guided dosing compares to standard dosing under each assumption. In the model, the “most likely,” “best,” and “worst” genotype-guided dosing scenarios are estimated to result in approximately 44, 40, and 52 serious bleeding events and 27, 24, 28, serious thromboembolic events, respectively.⁴⁷

Additionally, overall uncertainty related to data inputs can be evaluated using probabilistic sensitivity analysis, in which distributions are assigned to the model inputs, and Monte Carlo simulation is used to repeatedly draw sets of model inputs.^{51,52} Although the use of probabilistic sensitivity analysis is considered best practice for health outcomes modeling, use of individual parameter sensitivity analyses and multiple-parameter scenario analyses may be more intuitive for stakeholders.^{53,19}

Summary measure of health-related utility

Analogous to the USPSTF approach to presenting the results of indirect evidence assessments, we suggest presenting both benefits and risks in an outcomes table, as well as reporting ranges of results obtained from evaluations of uncertainty and assumptions, as described above.³⁵ However, assessing the overall balance of risks and benefits can be more challenging. Clinical events differ in their severity and frequency, and projecting their impact without an explicit framework or summary outcome measure is difficult. For example, considering the warfarin therapy cohort, should serious bleeding events experienced or thromboembolic events avoided receive a greater weight? Projected life expectancy is an important summary measure of mortality and should be assessed in all risk-benefit analyses for which there is uncertainty about clinical utility. However, life expectancy does not account for patient morbidity and quality of life impacts.

Quality-adjusted life-years

The challenge of comparing different types of outcomes across different diseases and interventions has been addressed in health outcomes research using the metric of the QALY.⁵⁴ The use of QALYs as the preferred measure in health outcomes research has been established in the United States and a variety of other countries.^{54,55} In addition, the recent Institute of Medicine Committee to Evaluate Measures of Health Benefits for Environmental Health and Safety Regulation in the United States stated that analyses that “integrate morbidity and mortality impacts in a single effectiveness measure should use the QALY to represent net health effects.”⁵⁶

The QALY represents an adjustment to length of life for the estimated quality of life. Quality of life is measured with a preference scale or index, where 0 represents the value or “utility score” for death and 1 represents normal “full” health.

Thus, 10 years of life expectancy at a utility of 0.5 is equivalent to 5 years with full health. There are several approaches to measuring preferences including time trade-off, standard gamble, and population-weighted surveys.⁵⁴ These measures evaluate physical, mental, emotional, and social functioning domains to varying extent and can be general or condition specific.

Grosse and Khoury suggested using the term “utility” to include both “clinical utility” (health-related outcomes) and “social utility” (primarily psychological effects).⁵⁷ We propose defining the utility of genomic testing from a health policy perspective as an improvement in life expectancy or quality of life for patients and their families, and term this measure “health-related utility” (HRU). The psychological impacts of testing, whether benefits or harms, would be included if they have a measurable impact on patient’s health-related quality of life, defined in general as mental, emotional, or social functioning related to their knowledge of genomic test results. In this construct, clinical events can be assessed through their impact on patient life expectancy (i.e., attributable mortality) and morbidity (i.e., patient quality of life). This definition combines attributes of “clinical utility” and “social utility” but does not include effects, such as impact on diagnostic thinking, if there is no associated influence on clinical outcomes or quality of life.

Returning to the warfarin case study, assessment of the potential impact of clinical events on life expectancy and QALYs requires the tracking of events, mortality, and quality of life over the lifetime of a patient cohort, which is commonly achieved in decision modeling through the use of Markov models.⁵⁸ We previously developed a warfarin pharmacogenomics health policy model using such techniques and estimated in the base-case analysis that testing could lead to an improvement in QALYs of 0.003 (1 day).⁴³ Notably, uncertainty analyses indicated that the difference in QALYs could range from -0.005 to $+0.010$. These findings are generally consistent with the results of similar analyses recently conducted by Eckman et al.⁵⁹ and Patrick et al.⁶⁰

Limitations of QALYs

There are several limitations to the use of QALYs as a summary measure of HRU for genomic tests. First, there are limited data on the impact of testing on patient and family quality of life or preferences.⁶¹ Second, measuring the psychological impacts of testing using a preference approach is challenging, because most instruments likely are not sensitive, and disease and test specific instruments will need to be developed.⁶¹ Third, there is significant uncertainty associated with most preference estimates, further complicating interpretation of the results of risk-benefit analyses. Fourth, different individuals will vary in their utility ratings of the same health state, so clinical guidelines should allow clinicians flexibility to address individual preferences, although population-level clinical policies will generally aim to consider average or typical preferences.

Many of these concerns have been noted by genomic test stakeholders in the literature.¹⁹ Specifically, we found that while stakeholders are receptive to the concept of using decision-analytic methods to evaluate genomic tests, many have concerns about lack of consistency in the methods used to elicit preferences, the ability of QALYs to capture the psychological value of test results, and the use of QALYs as a summary measure of HRU.¹⁹ Perhaps most importantly, stakeholders noted that use of QALYs as a summary measure of HRU is likely to lead to arguments about preference elicitation methods and could ultimately limit the use of decision-analysis to evaluate genomic tests.¹⁹

Table 1 Risk-benefit outcomes table: Examples of potential outcomes for different types of genomic tests

Incremental results for testing vs. comparator strategy	Type of genomic test		
	Pharmacogenomics	Disease risk	Newborn screening
Proportion of patients with a reclassified risk status	% patients with reclassified risk status	% individuals with reclassified risk status	% children with reclassified risk status
Proportion of patients indicated to receive an alternative treatment strategy	% patients with dose or drug change	% individuals initiating lifestyle change or preventative interventions	% children receiving medical or dietary intervention
Proportion of patients likely to choose the alternative treatment	% patients choosing indicated treatment strategy	% individuals choosing indicated lifestyle change or preventative intervention	% children receiving indicated medical or dietary intervention
Incidence of clinical events: benefits and harms (including NNT, NNS, NNB, NNH)	Serious adverse drug reaction, myocardial infarction, stroke, cancer recurrence	Myocardial infarction, stroke, cancer recurrence	Growth, mental functioning
Life expectancy	Life years	Life years	Life years
Quality of life (and patient preferences)	Impact of clinical events on quality of life	Impact of increased fear vs. empowerment	Impact of “right to know” vs. harm from uncertainty in diagnosis or treatment
Health-related utility	QALYs	QALYs	QALYs

NNT, number needed to treat; NNS, number needed to screen; NNB, number needed to benefit; NNH, number needed to harm; QALYs, quality-adjusted life-years.

These limitations highlight the need for ongoing stakeholder dialogue in relation to the development and use of decision-analytic methods to evaluate genomic tests, consideration of the impact of patient treatment preferences on health outcomes, and the importance of outcome measures in addition to QALYs. To help address these issues, we suggest that analyses report a multitude of health outcomes, including (1) proportion of patients with a reclassified risk status, (2) proportion of patients indicated to receive an alternative treatment strategy, (3) proportion of patients likely to choose the alternative treatment (4) clinical events (benefits and harms), (5) life expectancy, and (6) quality-adjusted life-expectancy (Table 1).

Risk-benefit policy matrix

Health policy evaluations of genomic tests are complex and warrant a variety of clinical, social, and political considerations. The framework established above serves to anchor one of these domains, HRU. Given the results of a quantitative risk-benefit analysis, in addition to other factors, decision-makers are faced with three options: (1) recommend the technology, (2) reject it, or (3) wait and collect more data.

In reference to the latter option, there has been increasing regulatory interest in the use of “coverage with evidence development” (CED). CED programs provide patients access to technology while developing evidence to inform future policy decisions.^{62,63} The U.S. Medicare program has recently applied this approach in other areas where there is limited evidence available (e.g., surgical interventions and medical devices). In such programs, health care payers agree to cover medical services or technologies under the condition that beneficiaries enroll in studies or registries to collect additional data on the use and outcomes of the therapy. Thus, CED provides a process for moving technologies along the translational pathway. For example, based on recommendation from the Medicare Evidence Development & Coverage Advisory Committee, Centers for

Medicare and Medicaid Services (CMS) recently implemented a CED policy for pharmacogenomic testing with warfarin therapy.⁶⁴

To implement CED in a manner consistent with facilitating the appropriate translation of genomics into health care, a “technology triage” mechanism is needed to identify potential candidates. We believe quantitative risk-benefit assessment can serve this important role. Risk-benefit policy matrices can be used to categorize genomic tests based on potential magnitude of HRU, and the uncertainty around these estimates. Our draft matrix (Fig. 3) provides five recommendation options, aiming to discourage use (or clinical development) of tests that have a reasonable chance of overall “negative” HRU, while encouraging entry into a “postmarketing” development pathway for tests that offer substantial promise but lack evidence of HRU.

For example, in the warfarin case study, model estimates indicated that genotype-guided dosing would result in a small increase in QALYs relative to standard dosing, but probabilistic sensitivity analyses estimated that genotype-guided dosing would increase QALYs in 84% of simulations and decrease QALYs in 16% of simulations.⁴³ Given these findings indicate an approximately “neutral” risk-benefit profile and a “moderate” degree of uncertainty, genotype-guided warfarin dosing could be classified as “use with evidence development.” The conclusion reached by the CMS in August of 2009.⁶⁵ Although formal decision modeling did not appear to have a direct role in this decision, CMS “considered the evidence in the hierarchical framework of Fryback and Thornbury where Level 2 addresses diagnostic accuracy, sensitivity, and specificity of the test; Level 3 focuses on whether the information produces change in the physician’s diagnostic thinking; Level 4 concerns the effect on the patient management plan and Level 5 measures the effect of the diagnostic information on patient outcomes.”⁶⁶ Although the evidence considered was similar, a formal risk-benefit approach may have provided greater transparency in regard to

Risk-Benefit Profile	Uncertainty		
	High	Moderate	Low
Favorable	Use with evidence-development	Consider use in clinical practice	Appropriate for use in clinical practice
Neutral	Do not use, conduct additional research	Use with evidence-development	Consider use in clinical practice
Unfavorable	Do not use, conduct additional research	Do not use	Do not use

Fig. 3. Risk-benefit policy matrix.

synthesis across levels of evidence, and quantification of the potential net benefit and associated uncertainty. Uncertainty analyses also could highlight the evidence gap in regard to the effectiveness of testing, and the value of conducting a randomized controlled trial.

The recommendation categories we propose for the risk-benefit framework offer a starting point for stakeholders to develop dialogue about the merits of genomic tests. As with most policy frameworks, we expect this approach to evolve over time and be modified as needed by individual stakeholder groups. For certain stakeholders, other considerations such as cost and equity will be important and should be evaluated. Indeed, in our previous study of stakeholder perspectives, payers indicated that tests with lower budget impacts might be evaluated using a simpler matrix primarily focused on potential harm, whereas more expensive tests or ones that have larger downstream cost impacts would require a careful evaluation of risk-benefit as well as cost-effectiveness.¹⁹

Challenges

In some cases, formal risk-benefit assessment of genomic tests will be limited by lack of sufficient or valid data to make utilization recommendations. In these cases, health outcomes modeling can be used to conduct exploratory evaluations to identify key parameter values that are required to produce HRU. For example, this approach was taken in the recent EGAPP evidence reports evaluating *CYP-P450* testing for antidepressant therapy and ovarian cancer susceptibility testing.^{39,40} In the *CYP-P450* for antidepressant therapy evaluation, a decision analysis was conducted to examine under which conditions genetic testing could lead to a better clinical outcome at 6 weeks, with the outcome of proportion responders.⁴⁰ In the ovarian cancer susceptibility testing evaluation, a Markov model with a lifetime horizon was used to assess what combinations of inputs would be required to achieve a target of 20% reduction in cancer mortality.³⁹

Quantitative evaluation of genomic tests is also complicated by their diverse applications,⁶⁷ and distinct ethical and policy implications, based on predictive value and the availability of treatment for patients who test positive.⁶⁸ In this sense, the risks and benefits of genomic testing extend beyond the usual endpoints measured in health technology assessment. Consideration of risks like stigma and discrimination, false reassurance, opportunity costs, and use of unproven therapies must be weighed against potential benefits of genomic diagnosis to family members, and the value placed on risk information by both patients and providers.^{57,68–72} Whether these risks and

benefits should drive health care decision-making is an open question, to be determined in part by their relative weight compared to medical outcomes of testing.

Finally, given the required assumptions and potential complexity of analyses, stakeholder acceptance of a modeling approach is likely to be a major challenge.^{72,73} To address this issue, collaboration with stakeholders to specify optimal approaches, interpretation, and recognition of limitations is critical to the success of a genomic testing risk-benefit framework. Such efforts are underway, but additional work is needed in this area.¹⁹ Issues that need to be addressed include (1) data to be included in risk-benefit analyses, (2) outcomes generated by the analyses, (3) a decision-making framework and corresponding thresholds, and (4) transparency, acceptance, and communication of the results of the analyses.

Summary

We believe a formal risk-benefit framework is useful for evaluating genomic tests for several reasons. First, although it must be recognized that the gold standard for direct evidence of HRU of genomic tests will come from prospective randomized controlled clinical trials, there is an opportunity to use quantitative risk-benefit analysis to derive at least preliminary estimates of HRU. This approach could be particularly valuable for genomic tests with a clear course of action that has been well studied.

Second, there will be a significant shortage of direct evidence of HRU for genomic tests in the near future. In some cases, indirect evidence of a favorable risk-benefit profile will suffice to recommend a test for use in clinical practice. Formal risk-benefit analysis offers a pragmatic approach to assessing HRU in a reasonably timely yet systematic manner. Thus, safe and potentially valuable genomic technologies will not be withheld from clinical use because of lack of direct evidence.

Third, risk-benefit analysis provides a tool to quantify the risk of interventions that result from testing in relation to potential benefit. Specifying risks can also aid in communicating such risks to providers and policymakers, thus protecting patients’ and the public’s health.

Finally, there will be significant uncertainty surrounding the HRU of most genomic tests. Scenario analysis and formal sensitivity analysis provide a mechanism for the quantification of uncertainty in HRU. Risk-benefit analyses also provide a foundation for assessing the value of additional research to reduce uncertainty and guide prioritization of comparative effectiveness research in genomics.

In summary, quantitative risk-benefit analysis provides a valuable tool for prioritizing genomic tests for development in the translational pathway. Specifically, tests that appear to have a reasonable risk profile, but with significant uncertainty with regard to the magnitude of benefit, can be recommended for use in clinical practice in CED programs. These strategies provide a viable route to generating evidence of HRU in the de facto “postmarketing” environment of genomic tests. This approach could serve as a foundation for assessment of population health impacts, regulatory decisions, health economics studies, and for the incorporation of the personal utility of prognostic information.

ACKNOWLEDGMENTS

Supported by Cooperative Agreement No. 1U18GD000005 from the Centers for Disease Control and Prevention (Veenstra D, PI).

REFERENCES

- Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: genetic testing strategies in newly diagnosed individuals with colorectal cancer aimed at reducing morbidity and mortality from Lynch syndrome in relatives. *Genet Med* 2009;11:35–41.
- Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: can UGT1A1 genotyping reduce morbidity and mortality in patients with metastatic colorectal cancer treated with irinotecan? *Genet Med* 2009;11:15–20.
- Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: can tumor gene expression profiling improve outcomes in patients with breast cancer? *Genet Med* 2009;11:66–73.
- Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: testing for cytochrome P450 polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. *Genet Med* 2007;9:819–825.
- Khoury MJ, McBride CM, Schully SD, et al; Centers for Disease Control and Prevention. The Scientific Foundation for personal genomics: recommendations from a National Institutes of Health-Centers for Disease Control and Prevention multidisciplinary workshop. *Genet Med* 2009;11:559–567.
- Khoury MJ, Feero WG, Reyes M, et al; GAPPNet Planning Group. The genomic applications in practice and prevention network. *Genet Med* 2009;11:488–494.
- Ramsey SD, Veenstra DL, Garrison LP Jr, et al. Toward evidence-based assessment for coverage and reimbursement of laboratory-based diagnostic and genetic tests. *Am J Manag Care* 2006;12:197–202.
- Garrison LP Jr, Carlson RJ, Carlson JJ, Kuszler PC, Meckley LM, Veenstra DL. A review of public policy issues in promoting the development and commercialization of pharmacogenomic applications: challenges and implications. *Drug Metab Rev* 2008;40:377–401.
- Garrison LP, Austin MJ. The economics of personalized medicine: a model of incentives for value creation and capture. *Drug Inf J* 2007;41:501–509.
- Hillman MA, Wilke RA, Yale SH, et al. A prospective, randomized pilot trial of model-based warfarin dose initiation using CYP2C9 genotype and clinical data. *Clin Med Res* 2005;3:137–145.
- Clarification of Optimal Anticoagulation Through Genetics (COAG), NCT00839657. Available at: <http://clinicaltrials.gov/ct2/show/NCT00839657>. Accessed July 19, 2010.
- Murphy GM Jr, Kremer C, Rodrigues HE, Schatzberg AF. Pharmacogenetics of antidepressant medication intolerance. *Am J Psychiatry* 2003;160:1830–1835.
- Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. *J Clin Oncol* 2008;26:721–728.
- Khoury MJ, Rich EC, Randhawa G, Teutsch SM, Niederhuber J. Comparative effectiveness research and genomic medicine: an evolving partnership for 21st century medicine. *Genet Med* 2009;11:707–711.
- Secretary's Advisory Committee on Genetics Health, and Society (SACGHS). U.S. system of oversight of genetic testing: a response to the charge of the secretary of HHS. Report of the Secretary's Advisory Committee on Genetics, Health, and Society, 2008. Available at: http://oba.od.nih.gov/oba/SACGHS/reports/SACGHS_oversight_report.pdf. Accessed July 19, 2010.
- Hudson KL, Murphy JA, Kaufman DJ, Javitt GH, Katsanis SH, Scott J. Oversight of US genetic testing laboratories. *Nat Biotechnol* 2006;24:1083–1090.
- Laje G, Paddock S, Manji H, et al. Genetic markers of suicidal ideation emerging during citalopram treatment of major depression. *Am J Psychiatry* 2007;164:1530–1538.
- Ray T. NeuroMark launches Suicidal Ideation Dx for investigational use pending final study. *Pharmacogenomics Reporter*. 2007.
- Roth JA, Garrison LP Jr, Burke W, Ramsey SD, Carlson R, Veenstra DL. Stakeholder perspectives on a risk-benefit framework for genetic testing. *Public Health Genomics*. 2010 Apr 20. [Epub ahead of print].
- Lesko LJ. The critical path of warfarin dosing: finding an optimal dosing strategy using pharmacogenetics. *Clin Pharmacol Ther* 2008;84:301–303.
- Garcia DA. Warfarin and pharmacogenomic testing: the case for restraint. *Clin Pharmacol Ther* 2008;84:303–305.
- Kangelaris KN, Bent S, Nussbaum RL, Garcia DA, Tice JA. Genetic testing before anticoagulation? A systematic review of pharmacogenetic dosing of warfarin. *J Gen Intern Med* 2009;24:656–664.
- Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G; American College of Chest Physicians. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest* 2008;133(6 Suppl):160S–198S.
- Garrison LP Jr, Towse A, Bresnahan BW. Assessing a structured, quantitative health outcomes approach to drug risk-benefit analysis. *Health Aff (Millwood)* 2007;26:684–695.
- Meltzer D. Risks and benefits of risk-management plans. *Health Aff (Millwood)* 2007;26:681–683.
- Pauly MV. Risks and benefits in health care: the view from economics. *Health Aff (Millwood)* 2007;26:653–662.
- Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Aff (Millwood)* 2007;26:741–748.
- Juhn P, Phillips A, Buto K. Balancing modern medical benefits and risks. *Health Aff (Millwood)* 2007;26:647–652.
- Cohen JT, Neumann PJ. What's more dangerous, your aspirin or your car? Thinking rationally about drug risks (and benefits). *Health Aff (Millwood)* 2007;26:636–646.
- Graham J, Hu J. The risk-benefit balance in the United States: who decides? *Health Aff (Millwood)* 2007;26:625–635.
- Califf RM. Defining the balance of risk and benefit in the era of genomics and proteomics. *Health Aff (Millwood)* 2004;23:77–87.
- Institute of Medicine of the National Academies. The future of drug safety: promoting and protecting the health of the public. Washington DC, IOM, 2006.
- Teutsch SM, Bradley LA, Palomaki GE, et al; EGAPP Working Group. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. *Genet Med* 2009;11:3–14.
- Calonge N. New USPSTF guidelines: integrating into clinical practice. US Preventive Services Task Force. *Am J Prev Med* 2001;20(3 Suppl):7–9.
- Sawaya GF, Guirguis-Blake J, LeFevre M, Harris R, Petitti D. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Ann Intern Med* 2007;147:871–875.
- Nelson HD, Huffman LH, Fu R, Harris EL. Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility: systematic evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2005;143:362–379.
- Centers for Disease Control and Prevention. Evaluation of Genomic Applications in Practice and Prevention (EGAPP): implementation and evaluation of a model approach. Atlanta, GA: CDC, 2006.
- Evaluation of Genomic Applications in Practice and Prevention (EGAPP). Working group: evidence reports. 2009. Available at: <http://www.egappreviews.org/workingrp/reports.htm>. Accessed July 19, 2010.
- Myers ER, Havrilesky LJ, Kulasingam SL, et al. Genomic tests for ovarian cancer detection and management. *Evid Rep Technol Assess (Full Rep)* 2006;(145):1–100.
- Matchar D, Thakur M, Grossman I, et al. Testing for cytochrome P450 polymorphisms in adults with non-psychotic depression treated with selective serotonin reuptake inhibitors (SSRIs). Evidence Report/Technology Assessment No. 146. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-02-0025.). Rockville, MD: Agency for Healthcare Research and Quality, 2007.
- Bonis P, Trikalinos T, Chung M, et al. Hereditary nonpolyposis colorectal cancer: diagnostic strategies and their implications. Evidence Report/Technology Assessment No. 150 (Prepared by Tufts-New England Medical Center Evidence-based Practice Center under Contract No. 290-02-0022) (AHRQ Publication No. 07-E008). Rockville, MD: Agency for Healthcare Research and Quality, 2007.
- Technology Evaluation Center (TEC). BlueCross BlueShield Association, 2008. Available at: <http://www.bcbs.com/bluesources/tec/>. Accessed July 19, 2010.
- Meckley LM, Wittkowsky AK, Rieder MJ, Rettie AE, Veenstra DL. A policy model to evaluate the benefits, risks and costs of warfarin pharmacogenomic testing. *Pharmacoeconomics* 2010;28:61–74.

44. Veenstra DL, Harris J, Gibson RL, Rosenfeld M, Burke W, Watts C. Pharmacogenomic testing to prevent aminoglycoside-induced hearing loss in cystic fibrosis patients: potential impact on clinical, patient, and economic outcomes. *Genet Med* 2007;9:695–704.
45. Garrison LP Jr, Carlson RJ, Carlson JJ, Kuszler PC, Meckley LM, Veenstra DL. A review of public policy issues in promoting the development and commercialization of pharmacogenomic applications: challenges and implications. *Drug Metab Rev* 2008;40:377–401.
46. Weinstein M, Fineberg H. Clinical decision analysis. Philadelphia, PA: Saunders, 1980.
47. Risk-benefit framework for genetic tests. University of Washington, Center for Genomics and Public Health, 2010. Available at: <http://depts.washington.edu/cgph/riskbenefit.htm>. Accessed July 19, 2010.
48. Garcia DA, Witt DM, Hylek E, et al; Anticoagulation Forum. Delivery of optimized anticoagulant therapy: consensus statement from the anticoagulation forum. *Ann Pharmacother* 2008;42:979–988.
49. Anderson JL, Horne BD, Stevens SM, et al; Couma-Gen Investigators. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. *Circulation* 2007;116:2563–2570.
50. van Walraven C, Oake N, Wells PS, Forster AJ. Burden of potentially avoidable anticoagulant-associated hemorrhagic and thromboembolic events in the elderly. *Chest* 2007;131:1508–1515.
51. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Med Decis Making* 1985;5:157–177.
52. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;17:479–500.
53. Weinstein MC, O'Brien B, Hornberger J, et al; ISPOR Task Force on Good Research Practices—Modeling Studies. Principles of good practice for decision analytic modeling in healthcare evaluation: report of the ISPOR task force on good research practices—modeling studies. *Value Health* 2003;6:9–17.
54. Gold ME, Russell LB, Siegel JE, Weinstein ME, eds. Cost-effectiveness in health and medicine. New York, NY: Oxford University Press, 1996.
55. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal (reference N0515). London: NICE, 2004.
56. Institute of Medicine of the National Academies. Evaluating measures of health benefits for environmental health, and safety regulation in the United States. Washington DC, IOM, 2006.
57. Grosse SD, Khoury MJ. What is the clinical utility of genetic testing? *Genet Med* 2006;8:448–450.
58. Briggs AH. New methods for analyzing cost-effectiveness. *BMJ* 2007;335:622–623.
59. Eckman MH, Rosand J, Greenberg SM, Gage BF. Cost-effectiveness of using pharmacogenetic information in warfarin dosing for patients with non-valvular atrial fibrillation. *Ann Intern Med* 2009;150:73–83.
60. Patrick AR, Avorn J, Choudhry NK. Cost-effectiveness of genotype-guided warfarin dosing for patients with atrial fibrillation. *Circ Cardiovasc Qual Outcomes* 2009;2:429–436.
61. Grosse SD, Wordsworth S, Payne K. Economic methods for valuing the outcomes of genetic testing: beyond cost-effectiveness analysis. *Genet Med* 2008;10:648–654.
62. Tunis SR, Pearson SD. Coverage options for promising technologies: Medicare's coverage with evidence development. *Health Aff (Millwood)* 2006;25:1218–1230.
63. Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payment to health outcomes: a taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers [published online ahead of print March 10, 2010]. *Health Policy*
64. Centers for Medicare and Medicaid Services. Proposed decision memo for pharmacogenomic testing for warfarin response (CAG-00400N). Washington, DC: CMS, 2009.
65. Jensen SK, Jacques LB, Ciccanti M, Long K, Eggleston L, Roche J. Decision memo for pharmacogenomic testing for warfarin response (CAG-00400N). Baltimore, MD: Centers for Medicare and Medicaid Services, 2009.
66. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11:88–94.
67. Burke W, Zimmern RL, Kroese M. Defining purpose: a key step in genetic test evaluation. *Genet Med* 2007;9:675–681.
68. Burke W, Pinsky LE, Press NA. Categorizing genetic tests to identify their ethical, legal, and social implications. *Am J Med Genet* 2001;106:233–240.
69. Johnson FR, Ozdemir S, Mansfield C, et al. Crohn's disease patients' risk-benefit preferences: serious adverse event risks versus treatment efficacy. *Gastroenterology* 2007;133:769–779.
70. Burke W, Atkins D, Gwinn M, et al. Genetic test evaluation: information needs of clinicians, policy makers, and the public. *Am J Epidemiol* 2002;156:311–318.
71. Burke W, Psaty BM. Personalized medicine in the era of genomics. *JAMA* 2007;298:1682–1684.
72. Clayton EW. Ethical, legal, and social implications of genomic medicine. *N Engl J Med* 2003;349:562–569.
73. Hughes DA, Bayoumi AM, Pirmohamed M. Current assessment of risk-benefit by regulators: is it time to introduce decision analyses? *Clin Pharmacol Ther* 2007;82:123–127.
74. Temple R. Quantitative decision analysis: a work in progress. *Clin Pharmacol Ther* 2007;82:127–130.