



Published in final edited form as:

Health Aff (Millwood). 2011 December ; 30(12): 2259–2268. doi:10.1377/hlthaff.2010.0637.

How Comparative Effectiveness Research Can Help To Advance ‘Personalized Medicine’ In Cancer Treatment

Scott D. Ramsey^a, David L. Veenstra^b, Sean R. Tunis^c, Louis P. Garrison^b, John J. Crowley^d, and Laurence H. Baker^e

^aResearch and Economic Assessment in Cancer and Healthcare (REACH) group, Fred Hutchinson Cancer Research Center, Seattle, WA

^bDepartment of Pharmacy, University of Washington, Seattle WA

^cCenter for Medical Technology Policy (CMTP). Baltimore, MD

^dCancer Research and Biostatistics (CRAB), Seattle, WA

^eGroup Chair’s Office, Southwest Oncology Group, University of Michigan Ann Arbor, MI

Abstract

The use of biomarkers to “personalize” cancer treatment —identifying discrete genes, proteins, or other indicators that can differentiate one type of cancer from another and enable the use of highly tailored therapies -- offers tremendous potential for improved outcomes and lower treatment costs. However, the rapid development of cancer biomarker, or genomic, tests— combined with a paucity of evidence to support the effectiveness of the tests—presents a challenge for patients, clinicians and other stakeholders. In this article, we propose that comparative effectiveness research be used to strengthen what is now a haphazard process for developing and marketing cancer biomarker tests. We suggest novel funding approaches and a systematic process for moving from regulatory approval to the generation of evidence that meets the needs of stakeholders and, ultimately, patients.

Keywords

personalized medicine; comparative effectiveness; genomics; cancer

Because of extraordinary technological advances in genomics and proteomics—the study of the genomes of organisms and the study of protein structures and functions, respectively—clinicians in cancer practice will soon have access to tens if not hundreds of biomarker tests, also known as genomic tests. Biomarkers are measurable variations between people in their genetic material, proteins, or other biological molecules. Clinicians can use biomarker test results that correlate with the presence of disease to make a diagnosis. Biomarkers can also be used to assess the risk of future disease, the aggressiveness of a patient’s disease over time, and the likelihood that a patient will respond to a particular treatment.

Although genomic tests offer the promise of effective and efficient cancer care tailored to the genetic profile of the patient or the type of cancer, many are being introduced into

Address for correspondence: Scott D. Ramsey, MD, PhD, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M3-B232, Seattle, WA 98109-1024. Phone (206) 667-7846, Fax (206) 667-5977, sramsey@fhcrc.org.

DISCLOSURE

The content of the article is solely the responsibility of the authors and does not necessarily reflect the views or policies of the National Cancer Institute.

clinical practice without sufficient evidence about their benefits, risks, and impact on health care costs, compared to standard care. David Ransohoff and Muin J. Khoury have argued that because personal genomic information can cause significant harms as well as benefits (for example, when a “positive” test results in patients starting a therapy that is not proven to help but has known side effects), it is no different from other kinds of clinical information and should be held to the same standards of evidence.¹

Comparative effectiveness research will be more important at the intersection of personalized medicine and cancer than anywhere else because of the high risks and costs of poor-quality test information about this group of diseases. We present an integrated approach to designing and implementing comparative clinical trials of personalized medicine technologies for cancer, such as the use of genomic tests. This process has been developed through a multidisciplinary collaboration between the Fred Hutchinson Cancer Research Center, the Center for Medical Technology Policy,² the University of Washington, and SWOG, a clinical trial cooperative.³

Key features of the approach are: (1) identifying high-priority personalized medicine technologies by reaching a consensus with multiple stakeholders, such as consumer groups and clinicians; (2) designing trials that meet the needs of multiple stakeholders; (3) conducting trials through the use of clinical trials groups, described below; and, (4) developing and implementing funding mechanisms shared by private and public insurers—such as Medicare’s policy of coverage with evidence development, which reimburses certain procedures if the patient’s data become part of a registry or are used in a clinical trial.

There are many potential benefits of using a process driven by multiple stakeholders to implement comparative effectiveness studies within a cooperative group setting such as clinical trials groups. These groups—which provide a standing infrastructure to conduct large-scale clinical trials have the capacity to generate high-quality evidence needed for several levels of decision making. Representatives of expert groups such as the National Comprehensive Cancer Network (NCCN) or the American Society of Clinical Oncology (ASCO) and insurers participating in a trial’s design phase can help shape studies so they are better able to inform practice guidelines and coverage policies. Finally, pooling research funds from several sources, such as the National Institutes of Health and insurers, enables the health system to evaluate a greater number of personalized medicine tests.

Problems in Developing and Evaluating Genomic Tests

Too Little Evidence of Effectiveness

A variety of personalized medicine applications in cancer treatment are in various stages of investigation, ranging from research to early clinical use. Although these technologies may improve care, providers and other stakeholders charged with assessing their usefulness in clinical practice have been hampered by the lack of information about these tests when they enter the market.

The Centers for Disease Control and Prevention’s Evaluation of Genomic Applications in Practice and Prevention group⁴ has conducted evidence reviews for genomic tests now on the market for colorectal, breast, and ovarian cancer. The group found that only half of the tests were consistently associated with clinical outcomes of interest such as disease recurrence or death, a concept known as clinical validity. There was reasonable evidence about roughly another quarter of the tests to suggest that they could measurably improve patient outcomes such as survival or quality of life or reduce the cost of care, compared to standard care without the test—a concept known as clinical utility.

Why do genomic tests often enter clinical practice with little evidence to support their clinical validity, utility, or cost-effectiveness? One important reason is the process by which the tests move from development through evaluation and regulatory review to the market. The search for an accurate biomarker often leads researchers to develop a test that identifies a particular risk or disease marker while overlooking the way that the risk or problem is currently assessed in practice. In addition, the process fails to consider what clinical actions may follow from the information provided by the test.

One example is the addition of genomic tests to standard clinical monitoring—including taking medical histories and performing regular physical examinations and mammograms—to assess breast cancer patients for the recurrence of disease after the initial treatment. The three biomarkers used in these tests—carcinoembryonic antigen (CEA) and two cancer antigens (CA), 15-3 and 27.29—are all antigens, a type of protein that tends to increase along with increases in the amount of cancer cells in the body. Studies suggest that monitoring these biomarkers may detect a recurrence of breast cancer five to six months earlier than standard monitoring. That earlier detection of recurrence might lead to timelier treatment, which could improve the patient's quality of life and chance of survival.

Unfortunately, these tests are neither highly sensitive nor specific for breast cancer recurrence—that is, they are neither very likely to show recurrence when it has truly recurred, nor likely to show definitively that recurrence hasn't happened. Results across tests of the same patient can fluctuate substantially, making it difficult to interpret changing levels. Because of their poor predictive ability, these tests could produce large numbers of false positive results, resulting in unwarranted anxiety on the part of patients and unnecessary additional evaluations such as computed tomography (CT) scans and biopsies. Moreover, the potential benefit of diagnosing a metastatic tumor five to six months earlier is unclear, because no cure for breast cancer at that stage exists today.

Because of these concerns—and the fact that no comparative effectiveness trials have been conducted specifically to evaluate whether tumor biomarker tests improve patients' survival rates or quality of life, or the cost-effectiveness of care—clinical guidelines do not recommend using these genomic tests to monitor patients for a recurrence of breast cancer after primary therapy, which is typically surgery.^{5,6} Despite these guidelines, our examination of records for 2007 from Medicare and the Surveillance, Epidemiology, and End Results program of the National Cancer Institute⁷ found that these genomic tests are ordered for more than one in ten women with breast cancer. Furthermore, the Centers for Medicare and Medicaid Services currently reimburses providers for performing the tests.

Too Much Regulatory Flexibility

The regulatory flexibility that manufacturers of genomic tests have in bringing their products to market is one reason for the lack of evidence about the tests.⁸

Commercial Test Kits—One pathway is through the development of a commercial test kit that manufacturers distribute to multiple laboratories. Makers of these kits must obtain Food and Drug Administration (FDA) approval prior to marketing them. The FDA requires evidence from human trials demonstrating that the tests are safe and effective when used for their intended purpose. For the FDA, *effectiveness* means that “it can fairly and responsibly be concluded by qualified experts that the device [test] will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling of the device.”⁹ The FDA is concerned about the possibility of false results, either negative or positive, rather than about whether a test represents value added, compared to standard care—its clinical utility.

Laboratory-Developed Tests—The second pathway to market is through the development of a test for a single laboratory. Such tests are commonly called laboratory-developed tests or, less formally, “home brew” tests. Traditionally these were low-risk diagnostic tests for well-defined conditions, specialized tests used to diagnose rare diseases, or tests designed to serve local patient populations. Examples of some laboratory-developed tests include microscopic examinations (such as Pap smears and manual cell counts), erythrocyte sedimentation rates, microbiology cultures and susceptibility tests, examination of tissue sections (including staining protocols), and blood cross-matching procedures. Tests for rare genetic disorders are typically designed for local populations.

With a few exceptions, laboratory-developed tests have not been subject to FDA review,¹⁰ although this policy is currently being reassessed.¹¹ Manufacturers of genomic tests have realized that choosing this route to market avoids the expense, time, and uncertainty that FDA review of commercial test kits entails.¹² As a result, the great majority of new genomic tests for cancer are brought to market this way.

Without criticizing the potential value of laboratory-developed tests, it is important to note that their manufacturers do not have to meet any federally specified evidentiary standard and can market the tests to clinicians and patients as they see fit. Jeffrey Shuren, director of the FDA’s Center for Devices and Radiological Health, testified that the FDA has observed problems with some laboratory-developed tests in recent years, including faulty data analysis, exaggerated clinical claims, fraudulent data, poor clinical study design, and unacceptable clinical performance.¹³

Certification of Tests—Most US manufacturers seek certification for their tests under the Clinical Laboratory Improvement Amendments of 1988, whether or not they have FDA approval.¹⁰ The Centers for Medicare and Medicaid Services regulates this program, setting quality control procedures to ensure that the tests that laboratories use are accurate and reliable. Separate and distinct from FDA approval, certification is not designed to evaluate the impact of tests on clinical practice or patient outcomes.

Public and private insurers have historically paid for laboratory tests that are certified by this program, according to Current Procedural Terminology (CPT) codes overseen by the American Medical Association.¹⁴ These codes are assigned to medical, surgical, and diagnostic services that are provided to patients and are used to determine the reimbursement that a provider will receive from insurers.

Two Tests for Breast Cancer

Gene expression profiling tests are genomic tests that identify the activity of genes in cancer cells in order to assess the patient’s risk of cancer recurrence. They are designed to provide better results than traditional risk assessments that rely on evaluating the physical characteristics of a tumor, such as its size and pathology.

Two gene expression profiling tests for breast cancer the—Oncotype DX Breast Cancer Assay and MammaPrint—are marketed to clinicians and patients throughout the United States. MammaPrint received FDA approval on February 6, 2007, for use in determining the likelihood that breast cancer would return within five to ten years after a woman’s initial cancer diagnosis and treatment.¹⁵ Oncotype DX reached the market in 2004 as a laboratory-developed test and thus did not receive FDA approval. On February 27, 2006 California’s Part B Medicare administrator agreed to cover Oncotype DX, manufactured by Genomic Health.¹⁶ Agendia, the manufacturer of MammaPrint, received approval from the administrator on November 26, 2009.¹⁷ Both manufacturers maintain labs in California.

The MammaPrint and Oncotype DX stories illustrate the uncertainties that manufacturers face about whether insurers will reimburse providers for using their tests, as well as the possible different approaches to development and marketing. Because each test costs approximately \$4,000, price was probably an important reason why major insurers waited for several years after the tests reached the market to cover them. Agendia's decision to seek FDA approval may have been based on a strategy to accelerate its acceptance by Medicare, commercial insurers, and possibly patients.¹⁸ However, the lengthy FDA approval process helped delay MammaPrint's entry into the market, compared to Oncotype's.

The flexibility that manufacturers of genomic tests have in pursuing or avoiding the FDA approval process contributes to the heterogeneity of evidence supporting the tests. This can be an important obstacle to evaluating the comparative effectiveness of different tests, even those with essentially identical uses. For example, Oncotype DX has evidence supporting both its ability to determine the likelihood of cancer recurrence (its prognostic validity) and to determine the potential benefit of chemotherapy in reducing the risk of recurrence (its predictive validity). MammaPrint has evidence only of prognostic validity.¹⁹ However, both tests are now covered by several state Medicare programs and most large private insurers.

An Innovative Framework for Developing Genomic Tests

As we have shown, the current development process for genomic tests is haphazard. We propose an innovative framework to replace it, which would meet the needs of multiple stakeholders, would be systematic and evidence-driven, and would encourage use of the right test at the right point in a patient's care. The framework includes systematic literature reviews and comparative effectiveness research.

Systematic Literature Review

Systematic literature reviews are an important first step in understanding the evidence for promising genomic tests in cancer. As an example, we used such a review, together with input from experts on the specific cancers to identify recent published and unpublished studies evaluating genomic tests for the five most common types of cancer: bladder, breast, colorectal, lung and prostate. We identified thirty-nine studies where the clinical validity of the test was explicitly evaluated (Exhibit 1). Only eight of the studies (21 percent) evaluated clinical validity using prospective correlation—that is, these studies were designed to evaluate the relationship between the result of a genomic test and the recurrence of cancer. None of the studies was a prospective evaluation of a test's clinical utility.

Selecting Tests for Comparative Effectiveness Research

If most genomic tests lack high-quality supporting evidence to start with, how should we select tests that warrant our spending time and money on more-informative prospective trials? We believe that the challenge of choosing tests for further study could be better addressed through early and sustained partnerships among experts and stakeholders, including test manufacturers, researchers, patients, consumers, clinicians, insurers, and members of groups that prepare guidelines for clinical practice and patients.

After conducting a systematic literature review like the one described above to identify tests that merit additional research, the next step involves sharing this information with stakeholders to find a "short list" of tests that have promising evidence and appeal strongly to the greatest number of groups. In the case of the tests in Exhibit 1, that short list might include those for common cancers or for conditions where the test provides uniquely valuable information that cannot be obtained by other means.

Each stakeholder group will have different perspectives and insights. Patient representatives will have perspectives about patients' values and preferences. Clinicians will understand the clinical context and the level of need for alternatives to the status quo. Test manufacturers and researchers will know what is currently being done to evaluate emerging genomic tests. Each of the groups might have insights into what groups of patients and points in the progress of cancer are particularly worth including in a new study.

Conducting Comparative Effectiveness Research on Genomics Tests

Comparative effectiveness research has been defined as “the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, monitor a clinical condition or to improve the delivery of care.”²⁰ Exhibit 2 compares features of comparative effectiveness research and traditional research on genomic tests.

Colorectal and Uterine Cancer

Some existing genomic tests have already been evaluated using study designs that could be considered comparative effectiveness research. An example is testing for Lynch syndrome. Caused by several genetic mutations, Lynch syndrome is a condition that predisposes patients to a high risk of developing certain cancers at an early age, primarily colorectal and uterine cancers. Although it was historically diagnosed through examining a patient's family history, there are now several types of genomic tests available to screen patients with newly diagnosed colorectal cancer for this condition. If a cancer patient is found to have Lynch syndrome, family members can be alerted, tested, and—if found to have the relevant mutations—given regular screening by colonoscopy and uterine swab.

A comparative effectiveness framework such as we propose would ask several questions about the merits of using these tests to identify patients with Lynch syndrome, compared to the alternative of using family history alone to determine the individual's likelihood of having Lynch syndrome.

The first question would be about the clinical validity of the tests: Do they have acceptable sensitivity and specificity -- in other words, are they likely to yield neither false negatives or false positives -- for identifying both people who have Lynch syndrome and those who do not? Summarizing available studies, the Centers for Disease Control and Prevention's Evaluation of Genomic Applications in Practice and Prevention group⁴ categorized the evidence supporting the tests as “adequate,” meaning that the studies are of reasonable quality and the tests have good sensitivity and specificity for the condition.

The second question relates to the clinical utility of testing: Is there sufficient evidence to recommend the use of genomics tests instead of using family history to identify people who should receive screening? The Evaluation of Genomic Applications in Practice and Prevention group reviewed available studies that addressed this clinical problem and again found “adequate” evidence in favor of the tests.²¹

For example, a study by Bonis PA, Trinkalinos TA, Chung M, et al. shows that first-degree relatives—parents, children, and siblings—of people with Lynch syndrome mutations are willing to undergo counseling (52 percent), gene mutation testing (95 percent), and to have colonoscopies every one to two years (80 percent).²² Finally, a large Finnish trial demonstrated that screening people with the mutations through regular colonoscopies reduced the incidence of colorectal cancer by 62 percent, compared to usual care.²³

The evidence supporting genomic testing of people without cancer for the Lynch syndrome mutations stems from retrospective and prospective cohort studies. Higher-quality studies—

specifically, controlled trials that are both prospective and randomized—may be necessary for tests that are directly tied to decisions about using chemotherapy. Much hinges on the outcome: A negative test result could trigger a decision to forgo chemotherapy in cases where the current standard of care is to receive it. That means that a false negative result might jeopardize a patient's chances for remission or cure, but the degree to which this is a problem remains unknown without a prospective, randomized study.

Breast Cancer

Oncotype DX and MammaPrint, the gene expression profiling tests for breast cancer discussed above in the regulatory context, are examples of genomic tests for which the evidence of comparative effectiveness is mixed. These tests measure different sets of genes but have the same purpose: to guide decisions about using adjuvant chemotherapy—that is, chemotherapy after a primary treatment such as surgery to—decrease the risk that cancer will recur. Despite their current use, there is no direct evidence supporting their clinical utility, compared to the traditional treatment of chemotherapy. In addition, no studies have directly compared the two tests to each other.

Patients who otherwise would have received chemotherapy but decide against it following genomic testing surely have a better quality of life because they avoid the side effects of chemotherapy, but they may suffer more anxiety about the recurrence of cancer, compared to those who are treated. In addition, insurers worry that women will demand the test simply to learn the likelihood that their cancer will recur but will choose to have chemotherapy regardless of the test result, rendering the \$4,000 test useless from a decision-making standpoint.

We have no information about what types of patients receive the tests and how well they adhere to treatment recommendations based on test results in actual clinical practice. We can answer those questions with retrospective studies, using health plans' medical records to determine the proportion of women who had the test and received chemotherapy.

Generating direct evidence of whether the tests lead to better outcomes, however, requires randomized clinical trials that compare patients who receive standard treatment to those whose treatment is based on the results of genomic tests. Two ongoing, prospective randomized trials will provide needed information: TAILORx for Oncotype DX and MINDACT for MammaPrint.

TAILORx randomly assigned nearly 10,000 women with early-stage, hormone-responsive, node-negative breast cancer and intermediate OncotypeDX Recurrence Scores—that is, scores of 11–25, which predict neither a high or low risk of cancer recurrence—to chemotherapy treatment followed by endocrine therapy versus endocrine treatment alone. Because more than 45 percent of Oncotype DX scores are in the range of 11–25, and because most women would receive chemotherapy using standard criteria, the study is primarily designed to provide information on the appropriate cutoff for recommending adjuvant chemotherapy. It does not directly assess the effect of the test result on clinical decision making. In contrast, MINDACT is a direct evaluation of MammaPrint's clinical utility, comparing the test to conventional methods of deciding whether a woman should receive chemotherapy.

Both trials include patient health outcomes such as survival and quality of life as endpoints, and both were started years after the test they are investigating reached the market.

One question for the future is whether high-quality trials of genomic tests can be designed before the tests become widely used in practice. Early trials are important because they can

help decision makers understand the risks and benefits of testing in a controlled setting, rather than through retrospective studies (with all the attendant problems of making inferences from observational data). Another question is whether, in the spirit of comparative effectiveness research, such studies can be designed to meet the needs of multiple stakeholders; that is, have designs that satisfy regulatory requirements, and provide information that is vital to other stakeholders, such as clinicians, patients, and health insurers. We believe they can, provided that they make use of existing clinical trials groups and explicit collaborations with insurers.

The Role of Clinical Trials Groups

Prospective randomized, controlled trials create the most compelling evidence to change clinical practice, but they are artificial constructs using highly selected patient populations. To address this issue, clinical trials groups could implement naturalistic or pragmatic trials. Pragmatic, or “large simple,” trials come closest to measuring the benefits of a novel treatment or approach in routine clinical practice. They minimize the criteria used to exclude patients and compare the innovation to usual care.

Purchasers and providers are requesting pragmatic trials to generate evidence supporting decisions about clinical care and health policy.²⁴ The Clinical Trials Cooperative Group Program supported by the National Cancer Institute offers an infrastructure for high-quality prospective comparative effectiveness research on genomics tests for cancer.²⁵ The program’s mission is to promote and support clinical trials of new cancer treatments and diagnostics, explore methods of cancer prevention and early detection, and study quality-of-life issues and rehabilitation during and after treatment.

To work most effectively with the Clinical Trials Cooperative Group Network we developed a multidisciplinary collaboration between the Fred Hutchinson Cancer Research Center, the Center for Medical Technology Policy,² the University of Washington, and SWOG³. The Center for Medical Technology Policy facilitates meetings with stakeholders—such as representatives from clinical practice, industry, patient advocacy groups, insurers, and other payers—to identify priorities for research among emerging genomic technologies.

Our first priority has been to design a prospective comparative effectiveness study for women enrolled in Medicare whose breast cancer has spread to their lymph nodes. The study will compare the usual treatment (chemotherapy for all patients) and treatment using results from Oncotype DX (chemotherapy only for patients whose recurrence risk scores suggest it would be appropriate). Stakeholders advised us on the criteria for patients’ admission to the study and identified endpoints to be investigated in the trial, which opened on January 15, 2011. We will track health care use and costs, quality of life, and patient preferences.

Paying For Better Evidence

Developing better evidence about the clinical utility of genomic tests for cancer will require creative funding strategies. The private sector bears much of the cost of discovering cancer biomarkers and manufacturing tests for them, but manufacturers cannot support definitive, long-term, randomized studies of health outcomes prior to obtaining reimbursement for their tests. Nor are federal sources of funding likely to be sufficient, given the current state of the US economy and the number of potential tests.

One policy mechanism that may be particularly useful is “coverage with evidence development.” As explained above, under coverage with evidence development, insurers reimburse providers for the cost of certain procedures if the patient’s data are added to a

registry or used in a clinical trial.²⁶ Medicare has used this approach for a variety of services, including off-label uses of biologics approved for colorectal cancer, the use of implantable cardioverter defibrillators to prevent sudden cardiac death, and the use of positron emission tomography to diagnose patients with malignancies or suspected dementia. One recent example is Medicare's relying on this approach to begin several clinical trials evaluating genetic tests for sensitivity to warfarin in patients undergoing long-term treatment to prevent coagulation.²⁷

Medicare's program was intended to provide financial support for promising technologies that are very costly, which many cancer genomic tests are, while further evidence about their effectiveness is obtained. In some recent cases, the costs of the genomic tests alone exceed the general research costs of the clinical trial. Thus, paying for tests used in trials has become a major concern for potential funding agencies.

Although Medicare's coverage with evidence development program has been in place for several years, the Center for Medical Technology Policy is investigating ways to make it feasible for private insurers instead to support evaluations of genomics tests.²⁸ SWOG and several major health plans are trying to resolve the related legal, financial, and policy issues.

In order for such an approach to be viable, clinical trials groups, insurers, and other interested stakeholders will need to develop standardized methods for assessing clinical validity and clinical utility, particularly in the context of rapidly evolving technology; establish a network of clinical organizations capable of rapidly enrolling patients in these studies; and develop collaborative mechanisms so that payers, clinicians, patients, and researchers can work together to design, fund, and implement studies.

Conclusion

Personalized medicine offers the promise of treatments tailored to individual patients, improved outcomes, and lower treatment costs for people with cancer. However, regulatory and market structures do not encourage the development and manufacture of genomic tests based on the needs of multiple stakeholders, nor do they promote the generation of high-quality evidence about the tests' role and value in clinical practice.

Comparative effectiveness studies can provide high-quality evidence that addresses multiple stakeholders' perspectives. But this research will require a clear process to translate research findings in the lab into clinical use in patients, as well as infrastructure—such as the Cancer Cooperative Groups Trials Network or the HMO Research Network (www.hmoresearchnetwork.org)[please provide]—and novel funding methods to pay for diagnostic tests or treatment interventions, including coverage with evidence development. Such an integrated model will require that test manufacturers, clinical trials groups, and insurers modify their current ways of operating and paying for trials and treatment, and will require flexibility and willingness to share risk if the model is to be successful.

Acknowledgments

The authors thank Karma Kreizenbeck and Judy Nelson for their assistance in preparing this article. This work was partially funded by the Center for Comparative Effectiveness Research in Cancer Genomics through the American Recovery and Reinvestment Act of 2009 by the National Cancer Institute, National Institutes of Health (Agency Award No. 5UC2CA148570-02).

References

1. Ransohoff DF, Khoury MJ. Personal genomics: information can be harmful. *Eur J Clin Invest*. 2010; 40(1):64–8. [PubMed: 20055897]

2. Center for Medical Technology Policy [home page on the Internet]. Baltimore (MD): Center for Medical Technology Policy; [cited 2011 Aug 1]. Available from: <http://cmtpnnet.org/>
3. SWOG [home page on the Internet]. Ann Arbor (MI): SWOG; [cited 2011 Aug 5]. Available from: <http://www.swog.org/>
4. Evaluation of Genomic Applications in Practice and Prevention (EGAPP) [Internet]. Atlanta, GA: Centers for Disease Control and Prevention; [updated 2011 Jul 29; cited 2011 Aug 5]. Available from: <http://www.cdc.gov/genomics/gtesting/EGAPP/>
5. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol*. 2007; 25(33):5287–312. [PubMed: 17954709]
6. Rojas MP, Telaro E, Russo A, Fossati R, Confalonieri C, Liberati A. Follow-up strategies for women treated for early breast cancer. *Cochrane Database Syst Rev*. 2000; (4):CD001768. [PubMed: 11034727]
7. National Cancer Institute. Surveillance, Epidemiology, and End Results [home page on the Internet]. Bethesda (MD): NCI; [cited 2011 Jul 29]. Available from: <http://seer.cancer.gov/>
8. Ferreira-Gonzalez A, Teutsch S, Williams MS, Au SM, FitzGerald KT, Miller PS, et al. US system of oversight for genetic testing: a report from the Secretary's Advisory Committee on Genetics, Health, and Society. *Per Med*. 2008; 5(5):521–8. [PubMed: 20490368]
9. Food and Drug Administration. Classification of Devices Intended for Human Use; Device Classes, 21 USC Sect. 360c [Internet]. FDA; [updated 2009 Apr 30; cited 2011 Aug 5]. Available from: <http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/FDCActChapterVDrugsandDevices/ucm110188.htm>
10. Centers for Medicare and Medicaid Services. Clinical Laboratory Improvement Amendments (CLIA): overview [Internet]. Baltimore (MD): Centers for Medicare & Medicaid Services; [updated 2011 Jul 12; cited 2011 Jul 29]. Available from: <http://www.cms.hhs.gov/clia/>
11. In 1992 the FDA asserted that all laboratory-developed tests were devices subject to regulation under the Federal Food, Drug, and Cosmetic Act of 1938; however, the agency has rarely exercised its enforcement discretion. Food and Drug Administration. FDA/CDRH public meeting: oversight of laboratory developed tests (LDTs), date July 19–20, 2010 [Internet]. Silver Spring (MD)FDA[updated 2010 Jul 16; cited 2011 Jul 29]. Available from: <http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm212830.htm#background>
12. Food and Drug Administration. Fed Regist 2001. Vol. 66. Silver Spring (MD): FDA; 2000. Agency information collection activities; submission for OMB review; comment request; medical devices; classification/reclassification; restricted devices: analyte specific reagents [Internet]; p. 1140-41.[cited 2011 Jul 29]. Available from: <http://www.fda.gov/OHRMS/DOCKETS/98fr/010501c.htm>
13. Shuren, J. Direct-to-consumer genetic testing and the consequences to the public: statement before the Subcommittee on Oversight and Investigations, Committee on Energy and Commerce, House of Representatives [Internet]. Silver Spring (MD): FDA; 2010 Jul 22. [cited 2011 Jul 29]. Available from: <http://www.fda.gov/NewsEvents/Testimony/ucm219925.htm>
14. American Medical Association. About CPT [Internet]. American Medical Association; [cited August 5, 2011]. Available from: <http://www.ama-assn.org/ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt/about-cpt.page?>
15. Food and Drug Administration. Press release, FDA clears breast cancer specific molecular prognostic test [Internet]. Silver Spring (MD): FDA; 2007 Feb 6. [updated 2009 Jun 18; cited 2011 Jul 29]. Available from: <http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2007/ucm108836.htm>
16. Centers for Medicare & Medicaid Services. LCD for Oncotype DX Test – Breast Cancer Prognosis (L20634) [Internet]. Centers for Medicare & Medicaid Services; [updated 2010 Mar 1; cited 2011 Aug 5]. Available from: http://coverage.cms.fu.com/mcd_archive/viewlcd.asp?lcd_id=20634&lcd_version=19&show=all
17. Centers for Medicare & Medicaid Services. Local Coverage Article for MAMMAPRINT Test-Breast Cancer Prognosis Coding Guidelines (A49426) [Internet]. Baltimore, MD: Centers for Medicare & Medicaid Services; [updated 2009 Sep 24; cited 2011 Aug 5]. Available from: <https://www.cms.gov/medicare-coverage-database/details/article-details.aspx?>

articleId=49426&ver=2&ContrId=174&ContrVer=1&CtrctrSelected=174*1&CoverageSelection=Local&ArticleType=All&PolicyType=Final&s=California++Northern&KeyWord=MammaPrint&KeyWordLookUp=Title&KeyWordSearchType=And&lcd_id=30376&lcd_version=5&show=all&bc=gAAAAABAAAA&

18. Agendia's website contains a "Everything You Want to Know About MammaPrint" section for patients, including the question, "Why is an FDA cleared test important to me?" Agendia. Everything You Want to Know About MammaPrint [Internet]. 2011. [cited 2011 Aug 5]. Available from: http://www.agendia.com/pages/questions___answers/142.php
19. 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(1):66–73. [PubMed: 19125125]
20. Sox HC. Defining comparative effectiveness research: the importance of getting it right. *Med Care.* 2010 Jun; 48(6 Suppl):S7–8. [PubMed: 20473202]
21. Palomaki GE, McClain MR, Melillo S, Hampel HL, Thibodeau SN. EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome. *Genet Med.* 2009; 11(1):42–65. [PubMed: 19125127]
22. Bonis PA, Trikalinos TA, Chung M, Chew P, Ip S, DeVine D, et al. Hereditary Nonpolyposis Colorectal Cancer: Diagnostic Strategies and Their Implications. *Evid Rep Technol Assess (Full Rep).* 2007 May.(150):1–180. [PubMed: 17764220]
23. Aaltonen LA, Salovaara R, Kristo P, Canzian F, Hemminki A, Peltomaki P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med.* 1998; 338(21):1481–87. [PubMed: 9593786]
24. Roland M, Torgerson DJ. What are pragmatic trials? *BMJ.* 1998; 24;316(7127):285.
25. National Cancer Institute. IOM Recommendations – NCI Clinical Trials Cooperative Group Program [Internet]. NCI; [cited 2011 Aug 5]. Available from: <http://transformingtrials.cancer.gov/initiatives/cooperative-groups>
26. Mohr PE, Tunis SR. Access with evidence development: the US experience. *Pharmacoeconomics.* 2010; 28(2):153–62. [PubMed: 20085391]
27. Centers for Medicare and Medicaid Services. Proposed decision memo for pharmacogenomic testing for warfarin response (CAG-00400N) [Internet]. Baltimore (MD): CMS; 2009 May 4. [cited 2011 Aug 1]. Available from: <http://www.cmms.hhs.gov/medicare-coverage-database/details/nca-proposed-decision-memo.aspx?NCAId=224&ver=15&NcaName=Pharmacogenomic+Testing+for+Warfarin+Response&NCDId=333&ncdver=1&IsPopup=y&bc=AAAAAAAIBAA&>
28. Kary, W.; Bergthold, L.; Talbott, M.; Aubry, W. Coverage for Evidence Development: A Conceptual Framework (Issue Brief) [Serial online]. Baltimore (MD): Center for Medical Technology; 2009 Jan. [cited 2011 Aug 1]. (Issue Brief). Available for download from: <http://www.cmpnet.org/cmtpr-research/applied-policy-and-methods/coverage-with-evidence-development>

Exhibit 1

Sources of Evidence in Thirty-Nine Recent Studies on Cancer Genomics

Type of cancer	Randomized controlled trial for test utility	Prospective correlation within a randomized controlled trial	Retrospective correlation within a randomized controlled trial	Observational study
Bladder	0	1	1	4
Breast	0	3	3	6
Colorectal	0	1	3	3
Lung	0	2	2	8
Prostate	0	1	0	1

SOURCE Authors' analysis of literature review.

NOTES Test utility is the ability of a test to provide useful clinical information about patient diagnosis and treatment of disease in order to improve health outcome. In a randomized controlled trial using prospective correlation, the study is initially designed to evaluate the relationship between the test result and the outcome of interest (in this case, recurrence of cancer). In a randomized controlled trial using retrospective correlation, after the study is completed, researchers determine the relationship between the test result and the observed outcomes.

Exhibit 2

Features of Comparative Effectiveness Research and Traditional Studies of Genomic Tests for Cancer

Feature of research	Comparative effectiveness research	Traditional studies
Priority of study among alternatives	As determined by multiple stakeholders, using criteria such as disease burden or cost, lack of information, variability in care	Clinical opportunity as dictated by expert assessment of emerging technology
Study design	Retrospective or prospective analysis	Retrospective analysis of existing tumor specimens; occasional prospective analysis of observational data
Comparisons	Direct comparisons of new therapy with usual care	Direct comparisons of competing therapies, often not considering usual care
Topics	Prevention, treatment, monitoring, and other broad topics	In most cases, prediction of narrow effects such as serious drug interactions, response to treatment, tumor recurrence
Perspectives	Multiple, including clinician, patient, purchaser, and policy maker	Clinician and patient
Study populations and settings	Representative of clinical practice	Highly selected
Data elements	Patient characteristics, quality of life, safety of treatment, resource use and costs, patient preferences	Patient characteristics, clinical endpoints
Funding	“Coverage with evidence development” programs, public-private partnerships	Private investors, research grants from federal sources such as NIH

SOURCE Authors' analysis.