Effective Health Care

Infrastructure To Monitor Utilization and Outcomes of Gene-Based Applications: An Assessment

Copyright Information:

Citation:

Infrastructure To Monitor Utilization and Outcomes of Gene-Based Applications: An Assessment

Investigators

Frank DeStefano, M.D., M.P.H.
Nedra Whitehead, Ph.D.
Linda J. Lux, M.P.H.
Kathleen N. Lohr, Ph.D.

Acknowledgments

Abstract

Background: With the completion of the human genome sequence, the development and utilization of gene-based tests are expected to proliferate. These tests may be used to help make early diagnosis, improve risk prediction, and target therapies for both traditional gene-based disorders as well as common chronic diseases. Thus, data are needed for public health surveillance of the utilization of gene-based tests to be able to monitor trends in use, appropriateness of use, and potential disparities in utilization. Health care policymakers, providers, and payers need data on how specific genetic tests and related interventions impact short- and long-term health outcomes, including information on cost-effectiveness. Information, however, is currently lacking on the use of gene-based tests and the outcomes of clinical interventions based on these tests.

Objectives: Our objectives were twofold: (1) to conduct an assessment of existing databases in the US health care system for monitoring the utilization and outcomes of gene-based applications (including tests and related interventions) in the health care system; and (2) to provide recommendations to establish appropriate and practical systems to assess use and outcomes of gene-based clinical applications.

Current Databases: The assessment included targeted reviews of the published and gray literature, discussions with key informants (some of whom formed a Technical Expert Panel [TEP] to provide additional guidance to the project), and information gathering about the capabilities of specific databases and surveillance systems. The project also included a small workshop in which experts and stakeholders discussed the initial findings of the assessment, gave their input on the strengths and weaknesses of different options and approaches for monitoring gene-based applications, and developed recommendations for future research and development.

Only limited, sporadic information is available on the utilization of gene-based tests over time. Some research and surveys suggest that knowledge on the part of some providers about the availability and utility of tests may be reasonably widespread and accurate. Little or nothing is known about the extent to which patients and their families are aware of tests and knowledgeable about their benefits and harms. Finally, there are few longitudinal data to indicate the benefits and risks of using genetic tests to guide interventions and medical decisions, such as in the selection of therapies, and their short- or long-term outcomes.

Recommendations and Future Directions: We identified no databases that could provide all the information desired about gene-based testing broadly. Current databases might be used alone or in combination to address specific questions about gene-based testing, although whether they would provide an adequate "monitoring" capability remains unclear. A number of challenges will have to be addressed before the ideal of being able to compile and link data from existing health databases and surveys for public health surveillance and health services research can be realized. These include developing standard codes for genetic tests and database architecture standards to allow interoperability between databases; addressing concerns about privacy and confidentiality; and reducing the proprietary and regulatory barriers that inhibit sharing of data. The needed technical advances, such as standards to facilitate database

interoperability, are beginning to be addressed by national initiatives to improve health information technology and promote personalized health care.

To further the development of databases for monitoring utilization and outcomes of gene-based applications, the following recommendations are made for additional research and development:

- Improve the coding of gene-based tests in many of the relevant databases so that the test type, reason for the test, and test results can be readily determined;

- Develop or adopt standards for the proper collection and storage of data from genetic testing laboratories for archiving the tests performed and facilitating interoperability between databases;

- Explore the possibility of adding questions to ongoing surveys or developing new surveys to monitor the availability of genetic testing centers, adequate counseling, and barriers to accessing counseling services;

- Consider establishing a survey of genetic testing laboratories similar to the National Ambulatory Medical Care Survey (NAMCS) for medical clinics and the National Hospital Discharge Survey (NHDS) for hospitals; and

- Develop pilot studies for a small set of diseases and tests.

Ultimately, the development and operation of databases and other data collection efforts to monitor utilization and outcomes of gene-based tests and related interventions will require public trust and support. This in turn will require confidence in the security built into the databases or information systems to protect personal health information and an understanding of the value of the databases in improving personal health care.

Key Words: Genomic, genetic, database, laboratory testing, utilization, pharmacogenomics, health outcomes, surveillance.

Contents

Tables

Appendixes

Chapter 1. Introduction

Background

The Secretary's Advisory Committee on Genetic Testing (SACGT) defined a genetic test as:

> *… an analysis performed on human DNA, RNA, genes, and/or chromosomes to detect heritable or acquired genotypes, mutations, phenotypes, or karyotypes that cause or are likely to cause a specific disease or condition. A genetic test is also the analysis of human proteins and certain metabolites, which are predominantly used to detect heritable or acquired genotypes, mutations, or phenotypes.*[1(p3)]

More than 1,000 gene-based tests are now clinically available, with an additional 300 available for research purposes only (e.g., GeneTests[2]). By 2009, the world market for gene-based testing is expected to reach $12.5 billion.[3] Much remains unknown about the effectiveness and appropriate use of these tests. Most are used for diagnosis of rare gene-based diseases, but a growing number have population-based applications, including carrier identification, predictive testing for inherited risk for common diseases, and pharmacogenetic testing for predicting drug response.

With the completion of the human genome sequence, new gene-based tests are expected to be rapidly developed; these are likely to include concurrent testing of multiple gene-based markers using microarray technologies (i.e., multiplex testing). These new tests may be used to screen populations or at-risk groups, help make early diagnosis, improve risk prediction, and target therapies for both traditional gene-based disorders and common chronic diseases. Conceivably, genetic tests might help explicate individual differences that account for disproportionate health outcomes or disparities among individuals and populations. These anticipated applications of genomic technologies have the potential for broad public health impact.

Understanding Gene-Based Tests and Interventions

Currently, little information is available on the validity and utility of gene-based tests. Analytical validity (the ability to measure genotype accurately and reliably) and, to a lesser extent, clinical validity (the ability to detect or predict the associated disorder or phenotype) are evaluated for some tests by the Food and Drug Administration (FDA) and the Centers for Medicare & Medicaid Services (CMS). FDA regulates the sale of kits for standardized tests under its mandate to regulate medical devices. CMS monitors laboratory practices and quality under the Clinical Laboratory Improvement Act (CLIA). Clinical utility (the net balance of risks and benefits associated with using a test in routine practice), however, has not been systematically evaluated for many gene-based tests. The Centers for Disease Control and Prevention (CDC) began a model project in 2004—Evaluation of Genomic Applications in Practice and Prevention (EGAPP)—to evaluate the evidence on validity and clinical utility of gene-based tests that are being transitioned from research to clinical and public health practice.

(More information about EGAPP can be found online at http://www.cdc.gov/genomics/
gTesting.htm.) The U.S. Preventive Services Task Force (USPSTF) also conducts clinical utility
reviews of selected gene-based tests.[4]

Although efforts such as EGAPP are under way to examine the validity and utility of gene-
based tests, the need to understand the use of specific tests and gene-based clinical interventions
and their short- or long-term outcomes is growing. The proliferation of new gene-based tests
makes it essential that clinicians or researchers develop criteria to distinguish tests that promote
health and prevent disease from those that have limited or no benefit or that may even pose harm
to the public. Examining whether gene-based tests are being used appropriately for screening
purposes or for diagnostic or predictive testing (e.g., based on early signs and symptoms of
disease or family history) is critical, as is determining whether laboratories and clinical settings
are exchanging relevant and useful information that promotes appropriate use and fosters
improved patient care. The Secretary's Advisory Committee on Genetics, Health and Society
(SACGHS) has endorsed the need for public health surveillance to assess the public health
impact of genetic testing, including appropriate utilization and health outcomes.[5]

Monitoring Use and Outcomes of Gene-Based Tests and Interventions

Despite these federal efforts to date, very little is known about the use of gene-based tests
and the outcomes of clinical interventions based on these tests. No system exists for monitoring
these tests after they become part of routine health care. No agency tracks the utilization of
genetic tests, whether tests are used appropriately, and the impact of the tests on clinical and
public health outcomes. Postmarketing surveillance and research are needed to provide
information on health care delivery, outcomes, and costs to support the translation of genetic
tests into clinical practice

As noted below, this project was concerned with understanding whether a monitoring system
is needed and, if so, of what it might consist. Monitoring, in this context, might be
conceptualized as a process or system for data acquisition, aggregation, and assessment. Among
the capacities one might expect such a system to have is the ability to link information from
disparate sources and to establish new ways to associate events, interventions, and outcomes.
This capability requires an infrastructure, appropriate methods to capture data, and the inferential
abilities to analyze the data to develop meaning and context. Generally, we use the term
"monitoring" to connote this entire spectrum of qualities and resources; some apply the term
"surveillance" to such activities, particularly in public health applications, but of course no
implication of investigation of individuals is intended.

Monitoring the integration of gene-based testing into health care requires a combination of
approaches because no single information source on gene-based testing exists. These approaches
include activities that might best be considered surveillance and others that fall more into a
"research" framework (particularly, in this context, applied research). In either case, creating
new databases, modifying existing databases, linking these databases, and making information
from them readily available to administrators, regulators, researchers, and others are all critical
steps. Thus, a coordinated system for monitoring gene-based testing might include analyses of

administrative and clinical databases (including diagnostic, pharmacy, hospital, and health plan databases), periodic surveys of health systems including clinics and laboratories, surveys of patient populations, or collaborative outcomes research carried out by a network of research centers.

With such a monitoring system for gene-based testing and applications, policymakers and clinicians could address numerous important public health issues. Three critical topics have high priority:

1. **Test Utilization and Knowledge of Test Results.** Utilization patterns include the number and type of tests performed, indications for testing, characteristics and practice setting of the providers ordering the tests, and basic demographics and health characteristics of the patients tested. Having access to reliable information of this type will allow public and private health care delivery systems to plan for increased demand for services related to genetic testing. It will also allow public health programs to identify activities that increase demand for testing, such as highly publicized cases and marketing campaigns. For example, one health maintenance organization found that referrals for genetic testing for breast cancer increased 244 percent during a direct-to-consumer marketing campaign.[6] In consumer and provider surveys, CDC also found increased consumer awareness of, interest in, and requests for testing following direct-to-consumer marketing, but population data were not available to assess testing trends.[7] With the introduction and proliferation of direct-to-consumer marketing of specific genetic tests and large genome-wide scans of hundreds of thousands of single nucleotide polymorphisms (SNPs), developing information on the validity and utility of such tests and making that information available to consumers and health care providers will be essential.[8] Monitoring the utilization of direct-to-consumer tests will require conduct of self-report surveys rather than reliance on health care databases.

   The hope is that genetic tests can help address individual differences in outcome that account for disproportionate differences or disparities among individuals and populations. Barriers to equitable access to gene-based tests and interventions, however, may arise in many circumstances. Physicians may lack access to information about the tests and their reliability, validity, and utility; be unclear when to order the tests and how to interpret them; or be resistant to using them. Patients may also fear gene-based tests and may be reluctant to agree to them. The EGAPP project is providing evidence-based information on gene-based tests. A monitoring system could address how well physicians are aware of and use the evidence-based information, and it could also address whether differences in the use of gene-based tests can be attributed to true differential access, patient choice, or other factors. For example, although several studies have shown disparities in the utilization of prenatal genetic screening tests, a recent study found few disparities in the proportion of women who reported that their health care providers discussed these tests with them.[9]

2. **Impact on Treatment Choices and Short-Term Clinical Outcomes.** Gene-based tests and interventions will profoundly affect, in predictable and unpredictable ways, the US health care system, providers and clinicians, and patients and their families. New opportunities for interventions will be identified; current preventive interventions or

treatments may be modified for some patients. A monitoring system to assess the impact of gene-based tests in guiding treatment decisions, the appropriateness of those decisions, and response to therapy (including adverse events) could provide essential information for clinicians and health care policymakers.

3. **Health and Long-Term Outcomes.** Ultimately, timely and accurate information on the impact of gene-based tests and interventions, if appropriately distributed and used, may help reduce mortality and morbidity, reduce health care costs, and improve quality of life among patients. With appropriate safeguards, this information will also allow public and private insurers to ensure that their clients have access to cost-effective tests and interventions.

The information needed to assess the utilization and impact of gene-based tests and interventions is both extensive and highly varied. It is also likely to be more sensitive to patients' and others' concerns about privacy, especially insofar as information from genetic tests for one person may have significant ramifications for his or her relatives.

Similarly, the sources of these types of information on use and outcomes of gene-based interventions will likely be diverse, reside in both the public and private sectors, and be variably accessible to decision makers and clinicians. Computerized databases of health care agencies; medical and genetic testing laboratories; and various state, regional, or national health surveys (e.g., National Health Interview Survey, Behavioral Risk Factor Surveillance Survey) may serve as the source for monitoring use of tests and certain interventions, particularly those based on lifestyle modifications (e.g., diet). Administrative and insurance claims data from both public and private insurers and payers may also play an important role in these monitoring efforts. The linkage of electronic medical records with genetic test results also affords opportunities for monitoring utilization and outcomes of gene-based applications. More specific data on provider and patient knowledge about gene-based tests and interventions, the appropriateness of their use, and the benefits and harms that they pose for patients may, however, require specially focused data collection efforts.

Aims of the Project

To foster progress in systematic acquisition and use of information for addressing the issues related to gene-based applications, the Agency for Healthcare Research and Quality (AHRQ), in collaboration with the CDC, commissioned the RTI International Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) Center to examine what would be needed to develop a monitoring system on genetic testing and related health care interventions in the US health care system. Clearly other forms of preventive and therapeutic strategies are critical to the levels of health and well-being in this country—and no comprehensive or single surveillance or monitoring system covers those health care interventions. Thus, in a broader policy context, readers cannot ignore issues such as the competing uses of health care resources, alternative health goals, and choices that must be made about what services to monitor for their impact on the quality and costs of health care. Nonetheless, compelling reasons exist to conduct an assessment specifically of gene-based tests and applications, including: (1) concerns about potential discrimination and the need for stringent privacy and confidentiality safeguards; (2) the implications of genetic test results for family members; (3) the increasing marketing of genetic

tests directly to consumers; and (4) the immense expectations that have been generated about the potential benefits for personal health care stemming from advances in genomics.

We were thus asked to concentrate on gene-based applications, which is the focus of this report. Specifically, the project had two primary goals: first, to assess existing systems and databases for monitoring the utilization and outcomes of gene-based applications in the health care system and, second, to provide recommendations to establish appropriate and practical systems to track use and outcomes of gene-based clinical applications.

To accomplish these goals, we conducted targeted reviews of the published and gray literature, held interviews with key informants in different health-related organizations, and gathered data about the capabilities of specific databases and surveillance systems. Consistent with the SACGT definition, we took a broad view of "genetic test" and included both germ line and somatic testing, as well as DNA-based tests and biochemical tests that have a genetic basis. Throughout our assessment, we sought guidance from a technical expert panel (TEP).

The core activity was a small workshop to bring together experts and a small number of well-defined stakeholders; they discussed the initial findings of the assessment presented in a preliminary report, gave their input on the strengths and weaknesses of different options and approaches for monitoring gene-based applications, and considered opportunities that may evolve from new initiatives such as the increasing use of electronic medical records and the development of data standards to facilitate data linkage and sharing. An important outcome of the assessment and workshop was guidance and recommendations that federal agencies, private health organizations, academia, and others can use to set priorities in establishing monitoring systems and research activities to ensure the safe and effective integration of gene-based clinical interventions into the health care system.

Following the workshop, we developed a draft report that was subjected to extensive eternal peer review. This final report is the ultimate product of the project.

Organization of This Report

In this report, we describe the current landscape of available administrative and clinical databases (such as diagnostic, pharmacy, hospital, health plan, and electronic health records) and compare the information available across different databases related to gene-based testing (e.g., rates of utilization of gene-based tests, indications for testing, subsequent treatment choices, outcomes). We also briefly assess the available health services research related to the use of gene-based testing and describe the gaps in what is known about the uptake and impact of gene-based testing and other gene-based interventions for the US population. We also summarize workshop deliberations.

Finally, from all these diverse sources, we develop an extensive set of recommendations for public agencies and private organizations to consider in establishing mechanisms for monitoring use and outcomes of gene-based interventions. Insofar as a comprehensive surveillance system (i.e., a single, national monitoring structure) is infeasible, we also note ways, based on rigorous research, to generate information useful for both policymaking and clinical decisionmaking.

Chapter 2 provides an overview of our assessment process and our findings on current available databases. In Chapter 3 we discuss the implications of our findings and make recommendations for future database development and research needs.

Appendix A provides details of our methods of data collection. Key informants are listed in Appendix B; Appendix C reproduces the key informants discussion guidelines-protocol. Appendices D and E give the workshop agenda and participants, respectively. Our literature search abstraction forms can be found in Appendix F. Finally, Appendix G lists additional databases that were suggested by our peer reviewers.

Chapter 2. Current Available Databases

The first step in the project was to conduct an assessment of current databases that might be used to monitor utilization, clinical impact, and outcomes of gene-based tests and related interventions. Data collection for the assessment comprised several related activities: reviewing both published and gray literature, conducting key informant interviews, reviewing genomics testing databases of several types, and, most critical, convening an invitational workshop to discuss issues relating to databases that could provide information on both use and outcomes of gene-based applications. This chapter briefly reviews our assessment methodology and then presents our findings on the assessment of current databases.

Overview of Assessment Process

In this section we provide a brief overview of our assessment methodology. A detailed description of our methods can be found in Appendix A.

Technical Expert Panel

At the beginning of the project, we convened a Technical Expert Panel (TEP) to guide the development of our assessment approach and methodology. The panel comprised experts from a variety of disciplines and organizations. TEP members provided consultation on the project methodology, helped identify additional experts to participate in the workshop, and provided comments on our final report. Three TEP members were also able to participate in the invitational workshop; several others served as peer reviewers of the draft final report.

Review of Peer-Reviewed Literature

We conducted a targeted literature review of the scientific literature that was intended to supplement both our gray literature search and our key informant interviews. These latter two data sources served as the primary methods of identifying relevant databases, but the scientific literature review provided some preliminary information. This activity was not meant to be a comprehensive review of the literature on the use or outcomes of gene-based tests. With more than 1,100 genetic tests in clinical use, and more than 15,000 publications on genetic screening, a complete review was beyond the scope of this project. Rather, the literature review was designed to identify:

- Existing information about the utilization and outcomes of gene-based tests and interventions;

- Databases and resources available and their strengths and limitations for assessing utilization and outcomes of gene-based applications; and

- Information regarding the optimal characteristics of a database or information network that could be used to assess the utilization and outcomes of gene-based tests and interventions.

Gray Literature Search

The sources for gray literature consisted of unpublished reports, news briefings, press conferences, web pages, and other information available electronically on the Internet. We appraised the documents identified through these searches for their relevance to the project. A majority of the documents addressed various aspects of genetic testing without referring to data or files containing data on genetic testing. In other cases, health information databases were cited without any specific reference to genetic testing. As a final step, we included only those documents containing information about data regarding genetic testing in this review.

Key Informant Interviews

RTI, with inputs from the Agency for Healthcare Research and Quality (AHRQ) and the Centers for Disease Control and Prevention (CDC), compiled an initial listing of individuals with knowledge of genetic databases. We included people who serve as data stewards or users of genetic testing data and completed interviews with 21 individuals (Appendix B). We sought representatives in each of five key areas: health services research, clinical laboratory management, genetics, industry (e.g., development of genetic tests), and federal agencies. We also received suggestions of other candidates who met our criteria from individuals who could not participate in the interviews but wanted to make sure that their organization or their expertise in the genetic database arena was represented. The goal of the interviews was to identify databases and mechanisms that might offer information about use of gene-based tests and interventions in health care in this country. Appendix C presents the semistructured interview protocols.

Laboratory Database Review

We also sought to understand the environment of genetic laboratories and their attendant databases, computer infrastructures, and related software programs. For this part of the assessment, we focused on molecular diagnostic laboratories to identify individuals who manage genetic testing data as part of their daily operations. We initially searched websites to identify these individuals. The next step was to identify key people who had knowledge of laboratory computer infrastructure and the strategic decisionmaking process, such as laboratory directors. Appendix B lists the individuals with whom we spoke specifically about laboratory databases.

Invitational Workshop

RTI organized the workshop, convened October 3-4, 2007, at the AHRQ Conference Center in Rockville, Maryland. The purpose was to elicit from participants guidance or possible recommendations for linking or modifying existing database systems or developing a new data system for monitoring and other uses set out by AHRQ. The final program, which included both plenary and breakout working sessions, appears in Appendix D. Three work group sessions were organized around the main themes of the project:

1. Test utilization and knowledge of test results

2. Impact on treatment choices and short-term clinical outcomes

3. Health and long-term outcomes.

Workshop participants included the following: TEP members; representatives of key stakeholder groups for this topic; experts on specific topics; key staff from AHRQ, CDC, the Food and Drug Administration (FDA), the Centers for Medicare & Medicaid Services (CMS), and other interested federal agencies, including the National Cancer Institute (NCI) and the Health Resources and Services Administration (HRSA); and RTI project staff. Appendix E shows the final list of participants.

We sent a preliminary draft of the current report to the participants to provide background information for the workshop deliberations. We asked workshop participants to review the initial findings in the preliminary draft report, identify strengths and weaknesses, discuss various possibilities for monitoring gene-based interventions, and consider opportunities that may be presented by developments in health information technologies, such as electronic medical records and standards for linking and sharing electronic health care data. They were also asked to propose a research agenda that would guide development of such systems and allow ongoing assessment of the use and outcomes of gene-based clinical interventions in the US health care system.

Much of the substantive work of the workshop was conducted in the three separate workgroups. To focus the workshop discussion, RTI staff provided a set of questions to the participants in a case study format. The selected case studies covered a range of issues: prevalence (common or rare), type of genetic variation (inherited or acquired), and clinical context (treatment or prevention), including two distinct examples of pharmacogenomics.

The examples were designed to focus the discussion on availability of information in the database on the following issues: (1) Patient population: asymptomatic persons or those diagnosed with a disease, (2) Test: type and results, and its utilization, (3) Impact on choice of intervention: screening/surveillance with another test such as mammography, surgery, or therapeutics, and (4) Impact on patient outcomes: mortality, morbidity. The selected case studies were intended to help identify to what extent information may be available from existing databases and to what extent we need to modify or create new databases and analytic tools.

Peer Review

After the workshop, the RTI Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) team revised and completed the draft report to take account of the workshop deliberations, conclusions, and recommendations. The report was then subjected to formal, external peer review by about 17 outside experts in the field. We revised the final report based on the comments and suggestions received in the peer-review process.

Findings on Current Databases

We present our findings in three sections. The first section provides a summary of the databases that were identified through the peer-reviewed and gray literature searches and the key informant interviews.

The second section provides more detailed findings on how identified databases have been used or could be used to address various questions according to the three broad purposes identified for our assessment. We follow the gene-based testing process from test awareness to health outcomes. For each step of the process, we discuss the existing information available on that step and the data sources from which that knowledge is derived; we also note other potential data sources. We describe each type of data source in detail the first time it is mentioned and then refer back to it when relevant, in subsequent steps.

The final section includes a summary assessment of current databases, including their limitations and their potential for contributing to a monitoring system for gene-based tests.

Identified Databases

Characteristics of Identified Databases. A summary of the characteristics and potential uses of databases identified in our assessment is provided in Table 1. The table presents the main categories of databases that will be described below and a few illustrative databases under each category. Key characteristics of each database are presented in the first section of table columns, including if the data are derived from a defined population, the geographic area covered by the database, the ability to link individual-level records in the database to medical records (such as an electronic medical records system), the ability to link records in the database to other public health databases (such as mortality records), and the availability of summary data. The second section of table columns (Information on Gene-based Testing Process Steps) summarizes the availability of data that would be relevant to various steps in the gene-based testing process as will be described below.

Sources of Identified Databases. To the extent possible, we relied on information published in the peer-reviewed literature, but much of the information on databases came from the gray literature and key informants.

Through the gray literature search, we identified one integrated electronic health record system, the Kaiser Permanente HealthConnect system that will eventually include the medical records of all Kaiser Permanente members. We also identified 18 laboratories that perform genetic tests. We identified them either directly from the Internet search or from the particularly comprehensive and informative website maintained by the University of Kansas Medical Center.[10] More detailed information was available from the websites of two laboratories that provide summary information on newborn screening: the National Newborn Screening and Genetics Resource Center[11] and the Genetics Home Reference.[12] Three large cancer databases were identified: the National Cancer Institute (NCI) Surveillance Epidemiology and End Results (SEER) registries, a network of population-based state and regional cancer registries covering 26 percent of the United States, which maintains a pooled database; the NCI Breast and Colon Cancer Family Registries, a multi-site research infrastructure of registries of families with high

Table1. Databases for monitoring gene-based tests

Database or Type of Database	Population based	Geographic coverage	Linkage to medical records	Linkage to public health data sets	Record or summary data available	Awareness	Screening criteria	Screening tests (Number)	Screening tests (Results)	Interpretation/followup	Diagnostic testing	Interpretation	Management	Family followup	Health outcomes
Public health databases															
Newborn screening information system	Y	National	N	Y	Y	N	Y	Y	Y	Y	Y	N	N	N	Y
Electronic medical records															
Kaiser-Permanente HealthConnect	Y	Partial state	Y	M	Y	N	M	Y	Y	Y	Y	Y	Y	Y	Y
Department of Veterans Affairs	N		Y	M	M	N	M	Y	Y	Y	Y	Y	Y	N	Y
Department of Defense	N		Y	M	M	N	M	Y	Y	Y	Y	Y	Y	N	Y
Clinical databases															
Genetics services clinics	N		Y	M	M	Y	Y	Y	Y	Y	Y	M	M	Y	Y
New York Regional Perinatal Systems	Y	Partial state	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N	N
Laboratory databases															
Quest Diagnostics	N		N		Some	N	N	Y	Y	N	Y	N	N	N	N
Labcorp	N		N		N	N	N	Y	Y	N	Y	N	N	N	N
Myriad Genetics	N		N		N	N	N	Y	Y	N	Y	N	N	N	N
Administrative databases															
CMS beneficiary claims data	Y	National	N	N	N	N	Y	Y	N	Y	Y	N	Y	N	M
Hospital discharge data	Y	State (census) or national (sample)	N	M	Y	N	M	N	M	N	N	N	N	N	N
Blue Cross/Blue Shield (FEHBP)	N		M	M	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Scheduled surveys															
Pregnancy Risk Assessment Monitoring System	Y	Multi-state	N	Y	Y	Y	N	Y	Y	N	Y	N	N	N	N
Medicare Health Outcomes Survey	Y	National	N	N	Y	Y	N	Y	Y	N	Y	N	N	N	N
National Survey on Ambulatory Medical Services	Y	National	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y
National Health Interview Survey	Y	National	N	Y	Y	Y	M	N	N	N	N	N	N	N	N

11

Table 1. Databases for monitoring gene-based tests (continued)

Database or Type of Database	Population based	Geographic Coverage	Linkage to medical records	Linkage to public health data sets	Record or summary data available	Awareness	Screening criteria	Screening tests Number	Screening tests Results	Interpretation/ followup	Diagnostic testing	Interpretation	Management	Family followup	Health outcomes
Behavioral Risk Factor Surveillance System	Y	National	N	N	Y	Y	M	N	N	N	N	N	N	N	N
Point-in-time surveys	Y	As needed	N	M	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ongoing research cohorts															
Marshfield Foundation	Y	Partial state	Y	Y	M	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Framingham Heart Study	Y	Town	Y	M	M	M	Y	Y	Y	Y	Y	Y	Y	Y	Y
Registries															
SEER	Y	Multi-state	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	N	Y
Other sources															
Kaiser Program on Genes, Environment and Health	Y	Partial state	Y	Y	M	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
National Breast and Colon Cancer Family Registries	N		Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Cancer Biomedical Informatics Grid	N		M	M	M	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Cancer Genetics Network	N		Y	Y	Y	M	Y	Y	Y	Y	Y	Y	Y	M	Y
Genomic Health /UHC OncotypeDX	N		N	N	N	N	N	N	N	N	Y	N	N	N	Y
Quintiles Informatics	N		N	Y	N	N	N	Y	N	N	Y	N	N	N	M
CAP Laboratory Accreditation Program	Y	National	N	N	N	N	N	Y	N	N	Y	N	N	N	N
Linked databases															
SEER-Medicare	Y	SEER catchment areas	M	M	Y	N	N	Y	Y	Y	Y	Y	Y	N	Y

CAP, College of American Pathologists; CMS, Centers for Medicare & Medicaid Services; FEHBP, Federal Employees Health Benefit Plan; M, Not clear if database has characteristic or can provide data on testing process step; N, Database does not have characteristic or cannot provide data on testing process step; SEER, Surveillance, Epidemiology, and End Results; UHC, United Healthcare; Y, Database has characteristic or can provide data on testing process step.

cancer risk, also with a pooled database; and the Women's Environment, Cancer and Radiation Epidemiology study of breast cancer. Basic information on which genetic tests are available and clinical and laboratory directories are available from the GeneTests website (http://www.genetests.org/). A future resource is the Kaiser Permanente Research Program on Genes, Environment and Health.[13]

Through the gray literature search, we also identified research resources that may be useful for special studies. The Human Genetics Initiative (http://www.nimh.nih.gov/research-funding/grants/nimh-human-genetics-initiative.shtml) has clinical or diagnostic information and immortalized cell lines for DNA extraction. The Autism Genetic Resources Exchange has biospecimens and clinical data from families with more than one case of autism spectrum disorder.[14,15]

Fifteen databases were identified during the key informant interviews (Table 2). These databases included electronic medical record systems, research databases including large research cohorts, laboratory performance system databases, newborn screening databases, registries, and insurance claims databases.

Table 2. Databases identified during key informant interviews

Database	Type of Database
Medicare and Surveillance, Epidemiology, and End Results (SEER)	Linked database: claims data linked to registry data
CanCORS	Research cohort
Cancer Research Network HMO Network	Clinical and administrative
Marshfield Clinic Database	Clinical
Marshfield Center for Human Genetics Database	Research database including clinical, laboratory and genotyping data
Kaiser Permanente HealthConnect™	Electronic medical record
National Health and Nutrition Examination Survey	Scheduled survey with physical examinations
Market Scan Databases	Vendor integrated clinical and claims database
caBIG™	Research network
CAP Accreditation and Lab Improvement database	Laboratory proficiency testing database
Newborn screening databases (several)	Public health databases with newborn screening results
Cancer Family Registry	Registry
Surveillance, Epidemiology, and End Results (SEER)	Registry
Blue Cross/Blue Shield Databases	Claims data
Federal Employees Health Benefit Plan	Claims data
Myriad Genetics	Laboratory data

caBIG, cancer bioinformatics grid; CanCORS, Cancer Care Outcomes Research and Surveillance Consortium; CAP, College of American Pathologists.

Our peer reviewers also suggested several databases that may have relevant information on gene-based applications. These suggestions were received after our formal assessment had been completed, but we list the suggested databases in Appendix G.

We may well have missed some relevant databases or other potentially useful sources of ongoing information about the use and outcomes of gene-based tests. We sought to identify possible databases through focused reviews of the peer-reviewed and gray literature and key informant interviews. The workshop also identified a few additional databases. Although the review of the literature was fairly comprehensive, it identified relatively few articles that described databases that could be used for monitoring utilization and impacts of gene-based tests. The gray literature review focused on identifying websites with relevant information; again, relatively little information was identified on databases of gene-based tests. Thus, much of the information that we were able to gather on potentially relevant databases came from key informants and workshop participants. We tried to identify key informants and workshop participants covering a broad range of expertise and with knowledge of a diverse set of databases.

Despite our efforts to be as comprehensive as possible in our search, we undoubtedly missed some relevant databases. Thus, our compilation of databases should not be viewed as exhaustive. For example, we focused on data from Kaiser Permanente as an example of databases from managed care organizations (MCOs) or integrated health care delivery systems. We are fully aware that similar data may be available from other MCOs, such as Group Health Cooperative in Seattle, or from other types of health plans, such as those of the family of services within the United HealthGroup umbrella. Similarly, Blue Cross Blue Shield is but one of many large insurance plans that may maintain potentially useful databases. Similarly, there are several commercial laboratory databases in addition to those we specifically mention. (A number of potentially relevant databases that were suggested by peer reviewers of this report are listed in Appendix G.) Nonetheless, we believe that we have fairly completely captured the various *categories* of databases. Although the listing may be incomplete, the specific databases that we identified can be considered exemplars of databases in each category.

Essentially no evidence base exists about the value of gene-based tests or interventions. Many, if not most, of the gene-based tests that are currently available have not been evaluated for their efficacy, effectiveness, or cost-effectiveness. We were concerned only with identifying databases that could be used to monitor utilization and impacts of the tests. We make no judgments on the utility of the tests, which is the focus of the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) program at the CDC.

Current and Potential Uses of Existing Databases

In this section, we present our findings on the existing information and data sources on gene-based testing and related interventions and outcomes, organized by the three main categories of our assessment and specific questions within those categories (Table 3).

14

For each step of the process, we list "questions of interest," which AHRQ had suggested in the initial request for the project, that a monitoring system might be able to answer. We then discuss the existing information available for that step and describe the data sources from which that knowledge is derived. We classify sources as:

- Public health databases

- Electronic medical records

- Clinical databases

- Laboratory databases

- Administrative databases

- Scheduled surveys

- Point-in-time surveys

- Ongoing research cohorts

- Registries

- Other sources, and

- Linked databases.

Each data source or type of data source is described in detail the first time it is mentioned, when relevant in subsequent steps, we refer readers back to those initial descriptions.

Table 3. Three assessment categories and related questions for the gene-based testing process

Test utilization and knowledge of test results

   Physician and patient awareness

   Screening criteria

   Screening tests

   Interpretation and followup of screening tests

   Diagnostic testing

   Interpretation of diagnostic tests

Impact on treatment choices and short-term clinical outcomes

   Appropriate clinical management of patients

   Appropriate family followup

Health and long-term outcomes

Test Utilization and Knowledge of Test Results

Physician and Patient Awareness

Questions of Interest. A monitoring system might answer the following questions related to physician and patient awareness:

- Do providers have access to information about the validity and utility of tests?

- Do patients have access to tests?

- Are there disparities in access?

15

Existing Information. Before a gene-based testing protocol can be implemented, either the physician or the patient must be aware of the test and interested in determining if it is applicable to the patient. We identified 12 articles that provided information on physician or patient awareness of one or more genetic tests.[16-27] The articles were published between 2000 and 2004 and reported on data from 1998 to 2001.

Mountcastle-Shah and Holtzman reported that 75 percent of primary care physicians (45 of 60) in the Washington, DC, area, who were interviewed in 1999, knew that carrier testing for cystic fibrosis was available; 90 percent knew that a test was available to identify an inherited susceptibility for breast cancer.[16] Batra et al. found that 34 percent of gastroenterologists in New York were aware that a test was available for hereditary nonpolyposis colorectal cancer (HNPCC); 52 percent were aware of the test for familial adenomatous polyposis.[17] Chase et al. reported that among physicians who treated patients enrolled in a genetic linkage study of Alzheimer's disease, 65 percent were aware of the test for apolipoprotein E (APOE, a diagnostic test for Alzheimer's disease in symptomatic individuals) and 37 percent were aware of the presenilin 1 test (a susceptibility or diagnostic test for use in families with multiple cases of early-onset Alzheimer's disease).[18]

Cho et al. reported that physicians who requested information or ordered a BRCA1/2 test at the University of Pennsylvania between October 1995 and October 1997 had learned about the test from journals (54 percent), professional meetings (47 percent), colleagues (29 percent), the lay media (11 percent), patients (4 percent), and other sources (10 percent).[19] Advertising also was found to increase awareness of gene-based tests. CDC found that, in 2003, both physician and patient awareness of BRCA1/2 testing was higher in Denver and Atlanta, where an advertising campaign had been conducted, than in Raleigh-Durham and Seattle, where no advertising campaign had taken place.[7] Physician awareness was 44 percent in Atlanta, 39 percent in Denver, 29 percent in Raleigh-Durham, and 18 percent in Seattle. Twenty-eight percent of physicians in Atlanta, but only 10 percent of physicians in Raleigh-Durham, reported that their patients had asked about testing.

One-third of physicians who treated Alzheimer's disease patients had received literature on APOE testing; slightly more than one-tenth had received literature on presenilin testing. Respectively, these figures represent approximately two-thirds of the physicians who had heard of APOE testing (65 percent) and one-third of those who had heard of presenilin testing.[18] Wideroff et al. reported that physicians who had received advertising for cancer susceptibility tests were twice as likely to have ordered a test as those who had not received advertising.[20]

Patient awareness is a major factor in the use of gene-based tests. Primary care practitioners in the Cincinnati region reported that patient interest was among the strongest motivations to refer patients for evaluation for hereditary breast cancer.[21] Forty-four percent of HMO-affiliated primary care clinicians in southeastern Pennsylvania and southern New Jersey reported that a patient had asked about cancer susceptibility testing. Among providers who reported that a patient had asked about testing, 70 percent actually ordered a cancer susceptibility test or referred a patient for testing; among those who had not been asked about testing by a patient, 11 percent ordered a test.[22]

After an advertising campaign for BRCA1/2 testing, physicians in Atlanta and Denver reported that patient requests for referrals and for testing increased and that they had ordered more BRCA1/2 tests.[7] Friedman et al. reported that 51 percent of primary care physicians in Texas had discussed cancer susceptibility testing when a patient initiated the inquiry, 51 percent had initiated a discussion of cancer susceptibility testing, and almost 30 percent had never discussed cancer susceptibility testing.[23] Wideroff et al. reported that physicians whose patients had asked about cancer susceptibility testing within the past year were more than 5 times as likely to report having requested a cancer susceptibility test.[20]

Fifty-four percent of US obstetricians and gynecologists reported that they offer carrier screening for cystic fibrosis if a patient requests it; 13 percent routinely offer carrier testing.[24] Among the 22 percent of US obstetricians and gynecologists who do not routinely discuss aneuploidy screening, 74 percent discuss it if the patient initiates discussion.[25] Among the one-third who do not discuss carrier screening for heritable disorders with all patients, 53 percent discussed it if the patient initiated discussions.[25]

Nearly two-thirds of obstetricians and gynecologists provided informational pamphlets on preconceptional screening, and more than one-third provided pamphlets on breast or ovarian cancer susceptibility testing.[26] Among New York gastroenterologists, 64 percent discussed genetic testing occasionally, 8 percent frequently, and 28 percent never.[17]

Data Sources. Existing information regarding physician and patient awareness of gene-based testing is drawn from surveys of physicians. Obtaining such data *routinely* through point-in-time surveys may be possible. Collecting such information would be difficult using any methodology other than surveys. Ongoing research may be another mechanism for extending knowledge about awareness and existing information.

Public Health Databases. Public health databases are compiled for a public health purpose, such as population screening or surveillance of one or more diseases, exposures, or risk factors. Some examples of public health databases are newborn screening databases, lead surveillance databases, and reportable disease surveillance files. We identified no public health databases with information on physician, population, or patient awareness of gene-based tests.

Electronic Medical Records. Electronic medical records may provide inferential data on physician awareness of gene-based tests, but they do not provide data on patient awareness of gene-based tests.

Clinical Databases. No clinical databases were identified that provide satisfactory information on physician or patient awareness.

Laboratory Databases. No laboratory databases were identified that provide information on physician or patient awareness.

Administrative Databases. Administrative databases do not provide information on physician or patient awareness.

Scheduled Surveys. Two articles used data from the NCI Physician Survey on Cancer Susceptibility Testing, a nationally representative survey of physicians in selected specialties.

NCI conducted the survey, which questioned physicians about their practices and attitudes regarding cancer susceptibility testing, in 1996 and 2001.[20,27] This survey could be incorporated into an ongoing surveillance system on gene-based tests to provide information on physicians' knowledge of gene-based tests if NCI plans to repeat it periodically.

The National Center for Health Statistics (NCHS) conducts two annual surveys of outpatient clinics: the National Ambulatory Medical Care Survey (NAMCS) and the National Hospital Ambulatory Medical Care Survey (NHAMCS). The NAMCS is a national survey of physicians that yields information about the provision and use of ambulatory medical care services based on a sample of patient visits to physicians engaged primarily in direct patient care. The current instrument does not collect any information on gene-based testing, but the section on testing could be expanded, or a gene-based testing supplement, as is currently being done for a cervical cancer screening supplement, could be developed. NAMCS public use data files are available.[28] NHAMCS collects data on ambulatory care services provided through hospital emergency and outpatient departments. The NHAMCS instruments do not collect information on gene-based testing, but the instruments could be expanded or a supplement could be developed.[29]

Using existing national health surveys to collect information on the general awareness and attitudes about gene-based testing among the general population and some subpopulations is possible, although few surveys currently collect such information. The National Health Interview Survey (NHIS) is a cross-sectional household interview survey of approximately 43,000 households, including about 106,000 persons, conducted annually. NHIS monitors the health of the US population on a broad range of health topics.[30]

The National Health and Nutrition Examination Survey (NHANES) provides much broader information than NHIS. The NHANES detailed interview includes demographic, socioeconomic, dietary, and health-related questions. The examination component consists of medical and dental examinations, physiological measurements, and laboratory tests administered by highly trained medical personnel. The survey has become continuous, with a changing focus that addresses emerging health topics. The sample size, however, is limited to about 5,000 persons each year.[31]

The Behavioral Risk Factor Surveillance System (BRFSS) is an on-going survey conducted by state and territorial health departments that produces more than 350,000 interviews per year. The BRFSS tracks health conditions such as asthma, diabetes, and hypertension and risk behaviors such as health care access, alcohol and tobacco use, obesity, cancer screening, and nutrition and physical activity.[32] The Pregnancy Risk Assessment Monitoring System (PRAMS), an ongoing survey of women with a recent live birth, could be a source of data on patient knowledge and awareness of prenatal or neonatal gene-based tests.[9]

Point-in-Time Surveys. The American College of Obstetricians and Gynecologists (ACOG) maintains the Collaborative Ambulatory Research Network, a volunteer sample of obstetricians and gynecologists who are willing to be surveyed on knowledge and practices; three of the studies that we identified on this topic used this source.[24-26] The survey panel was designed to be representative of the population of US obstetricians and gynecologists and it may be a resource for relatively quick surveys of gene-based testing knowledge or attitudes in the obstetrical and gynecological communities.

Ongoing Research Cohorts. Ongoing longitudinal prospective studies or research collaboration could provide data on several aspects of gene-based testing. An example of such a study is the Framingham Heart Study, which has studied cardiovascular disease since 1948. The study follows participants over a long period of time and collects detailed information from participants, including medical history, physical examination, and laboratory tests. This study may contain information related to awareness of gene-based testing, or if not, could collect such information with the researchers' collaboration. The study is currently recruiting a Generation III cohort.

Registries. Some registries may serve as a sampling frame for surveys of patient awareness, although they do not routinely collect such information.

Other Sources. The Cancer Research Network consists of research programs, enrolled populations, and data systems of 12 health maintenance organizations (HMOs) nationwide. As part of its early detection cancer research programs, the Network has the potential to collect information on patient and physician awareness regarding gene-based testing.

Linked Databases. No databases were identified that could be linked to provide information on physician or patient awareness.

Screening Criteria

Questions of Interest. A monitoring system might answer the following question related to screening criteria: Are patients referred appropriately for gene-based testing?

Existing Information. For this report, we defined screening criteria as a questionnaire or other clinical tool that determines who should be referred for screening or diagnostic gene-based testing. Screening criteria for many conditions are based on family history information, although tests may also apply to entire segments of the population (such as pregnant women) or be based on disease symptoms.

Many physicians do not collect the information that they would need to assess whether patients meet the criteria for relevant gene-based tests. Wilkins-Haug found that 24 percent of US obstetricians and gynecologists reported that they did no assessment of family history or DNA-based testing.[26] Morgan et al. found that only 48 percent routinely asked their nonpregnant patients about family history, although an additional 35 percent asked about family history if the women were trying to get pregnant.[24] Although 89 percent of obstetricians and gynecologists ask their pregnant patients about a family history of cystic fibrosis, only 36 percent asked women who were trying to get pregnant about a family history of cystic fibrosis.[24] Only 13 percent routinely asked their nonpregnant patients about a family history of cystic fibrosis.

Summerton and Garwood found that although the majority of the British physicians they surveyed felt that family history was important in decisionmaking in several clinical scenarios, fewer than half collected family history information on any of the specified conditions except coronary heart disease.[33] Screening criteria for cancer susceptibility testing can be guided by cancer history, including the age at diagnosis, in first-, second-, and third-degree relatives. Almost all primary care physicians and gastroenterologists surveyed on this issue reported that

they collect data on first-degree relatives (90 percent and 99 percent, respectively), but only 56 percent of primary care providers and 76 percent of gastroenterologists collected data on second-degree relatives.[17,22] Only 39 percent of gastroenterologists collected data for third-degree relatives.[17]

The simplest screening criteria are those that apply to an entire population, such as newborn or prenatal screening. For more than 10 years, clinical practice guidelines have recommended that all pregnant women be offered screening for neural tube defects and Down syndrome.[34] In 2004, only 78 percent of US obstetricians and gynecologists routinely discussed aneuploidy screening with their patients.[25] A higher percentage discussed aneuploidy screening with women with an identified risk factor, such as advanced maternal age (92 percent), a significant medical or family history (85 percent), or an abnormal ultrasound (80 percent). Eighty-six percent of women in a large multistate study who began care in the first trimester reported receiving information on genetic screening.[9] Women under 18, Spanish-speaking women, and women whose prenatal care was not covered by insurance were less likely to have received information on genetic screening than other women. Among women in central New York who received care in hospital or community clinic settings, African-American women were less likely than white women to be offered a maternal serum-alpha-fetoprotein (MSAFP) test.[35]

In 1997, a National Institutes of Health (NIH) Consensus Development Conference recommended that carrier screening for cystic fibrosis be offered to all couples planning a pregnancy or seeking prenatal care, to adults with a family history of cystic fibrosis, and to partners of persons with cystic fibrosis.[12] In 2004, only 67 percent of US obstetricians and gynecologists reported that they discussed carrier screening for heritable diseases with all their patients.[25] Eighty-eight percent discussed carrier testing if a woman had a family history of a heritable disease, but only 58 percent discussed it with women whose partners had a heritable disease. According to Morgan et al., 13 percent of US obstetricians and gynecologists routinely offer cystic fibrosis screening to their nonpregnant patients, 18 percent would offer testing in all scenarios covered by the screening guidelines, but 19 percent never offer cystic fibrosis screening to their patients. [24] Almost two-thirds of obstetricians and gynecologists offer cystic fibrosis carrier screening to all their pregnant patients, and only 2.2 percent said they never offer cystic fibrosis carrier screening. Of those who offered testing in some scenarios, however, only 27 percent offered screening in all scenarios covered by the screening guidelines.[24]

Carrier screening may also be performed when it is not recommended. One national testing laboratory reported a 30-fold increase in testing for Fragile X syndrome between 1992 and 2006, with most of the increase occurring during the most recent years because of increased testing of adult women.[36] Very few males or females were tested because of a positive family history, and the authors concluded that the increase could be attributed to population-based carrier screening, which is not currently recommended by any professional organization.[36]

Gene-based testing screening criteria for conditions that can be either familial or sporadic, such as breast or colorectal cancer or Alzheimer's disease are more complex. The Amsterdam I and II criteria are used to identify families that should have gene sequencing performed to identify hereditary nonpolyposis colorectal cancer (HNPCC) mutations in the DNA mismatch repair genes.[37,38] The Bethesda Guidelines identify which tumors with colorectal cancer should be tested for microsatellite instability, a marker of HNPCC colorectal tumors.[37] Kievit et al.

conducted a meta-analysis of the sensitivity (the probability that a person with the condition of interest is identified by the screening test or guidelines) and predictive value positive (the probability that a person with a positive screening test or who meets the guideline criteria has the condition of interest) of these criteria. They found considerable heterogeneity among articles in both the sensitivity and predictive value of the Amsterdam criteria I with regard to identifying mutations; sensitivity ranged from 54 percent to 91 percent, and predictive value positive ranged from 62 percent to 84 percent. For the Amsterdam criteria II with regard to identifying mutations, sensitivity was 78 percent and positive predictive value ranged from 46 percent to 68 percent. The Bethesda guidelines, which identify patients whose tumors exhibit microsatellite instability as probable cases of HNPCC, had a sensitivity of 89 percent and predictive value positive of 53 percent.[37] A more complete description of the performance of various types of screening for HNPCC is available from a recent evidence-based review.[39]

Lynch et al. illustrated in several case scenarios the difficulty of applying these criteria in clinical practice.[38] Rarer mutations in mutation repair genes can result in atypical HNPCC presentation, or atypical presentation of other familial cancer syndromes can result in HNPCC phenocopies.[38] Physicians may fail to apply the screening criteria appropriately, either because they are not knowledgeable about genetic cancer syndromes or because they collect insufficient family history.[38] Seventy-nine percent of gastroenterologists could recognize pedigree as strongly suggestive of HNPCC.[17] Almost 91 percent of physicians who were aware of APOE testing for Alzheimer's disease understood that it should not be used for testing asymptomatic individuals. However, only 39 percent of physicians who were aware of the presenilin 1 test understood that it is an appropriate screening test for asymptomatic individuals who are under 50 years of age and have many family members with early-onset Alzheimer's disease.[18]

Data Sources. Existing information on physicians' knowledge and application of screening criteria for gene-based testing is based on data from clinical and laboratory databases and physician surveys.

Public Health Databases. Birth defects surveillance program data and vital statistics data could be combined to provide information on the appropriate application of prenatal screening. The Central New York Regional Perinatal Data System (RPDS), which is a population-based birth registry that includes information on demographics, antepartum, intrapartum, and postpartum care, and neonatal outcomes, has been used for this purpose.[35] Similar data systems are available for other regions of New York. The RPDS and similar databases could be used to monitor utilization and outcomes of pregnancy-associated gene-based tests.

Electronic Medical Records. Linked electronic medical records are a potential source of data for almost all aspects of gene-based testing, including adherence to screening criteria. As one example, Kaiser Permanente's electronic health record system, KP HealthConnect™, is a comprehensive health information system that includes one of the most advanced electronic health records available in the United States. When fully deployed, the data system will have information on 8.6 million people; it will capture data from the physician's office, the hospital, the radiology department, the laboratory, and the pharmacy. It is an integrated data system; similar data are maintained in all facilities within a region and then synchronized across the region. Its primary purpose is to improve the quality of care and service to KP members. The database could provide extensive information on genetic testing services for Kaiser's patients.

21

Kaiser Permanente of Northern California has five genetic centers with medical geneticists with subspecialty training, molecular and cytogenetics laboratories, genetic counselors, genetic nurses, and metabolic nutritionists. The program includes prenatal screening and diagnostic services, neonatal screening, multispecialty clinics for common genetic disorders, and adult genetic services, including cancer genetics, clinical genetics, and carrier screening.[13]

Clinical Databases. Lynch et al. drew their case studies' clinical records from their own practice and those of colleagues in a neighboring state.[38] The studies included in the meta-analysis by Kievit et al. were based on clinical records and research studies.[37]

NCI maintains the Cancer Research Network (CRN), a network of 12 large health maintenance organizations (HMOs) that use the group or staff model of service provision, with combined memberships of 11 million people. The HMOs provide annual clinical and administrative database updates. Currently 8 or 9 years of data are available. The overall goal of the CRN is to conduct research on cancer prevention, early detection, treatment, long-term care, and surveillance. Data elements commonly used on multiple CRN projects are identified and prioritized through a data mapping initiative. Conceptual and operational definitions of the variables are established; from these definitions, analysts write computer programs that can be used at all sites to extract data. Ultimately a library of programs and macros will be created. Through this effort, CRN sites are also able to identify limitations to data, including quality and availability, and other idiosyncrasies (personal communication, Arnold Potosky, Ph.D., NIH, Health Services and Economics Branch, Applied Research Program, July 31, 2007).

Laboratory Databases. Laboratory databases may include information on family history or the reason for the test, although at least one laboratory, Quest Diagnostics, finds that the indication for testing is rarely provided.[36] Information on the testing indication and the appropriateness of testing can sometimes be inferred from other information in the database, as in the discussion of changing trends in Fragile X testing by Strom et al.[36] The potential role of laboratory databases as a component of a surveillance system on gene-based tests is discussed in more detail in the next section.

Administrative Databases. Administrative databases are usually not specific enough to provide information on the correct application of screening criteria. Some databases, however, may contain codes for genetic counseling or the collection of family history.

Scheduled Surveys. PRAMS, an ongoing survey of women with a recent live birth, includes questions on the provision of information regarding genetic screening during prenatal care.[9] Questions on sickle cell testing are being considered for inclusion on the new questionnaire. (personal communication, Tonya Stancil, Ph.D., DC, August 18, 2007). This survey accepts question submissions from the public health community and may be a source of data about prenatal or neonatal gene-based tests.

Point-in-Time Surveys. Most of the existing information on the use of screening criteria and physicians' understanding of risk are based on point-in-time surveys that are conducted for a specific research project and that examine screening criteria for a specific condition. The populations surveyed are frequently limited to a small or specific group of physicians.

Ongoing Research Cohorts. Clinical data sets that are supplemented with genetic information for research purposes may also provide useful data. One large HMO currently maintains such a database through its research foundation; a second is designing a similar database.

The Center for Human Genetics at the Marshfield Clinic Research Foundation maintains the Personalized Medicine Research Project, a population-based biobank with nearly 20,000 subjects ages 18 years and older, access to electronic medical records to classify phenotype, as well as DNA, plasma, and serum samples. The Kaiser Permanente Research Program on Genes, Environment and Health is designed to identify genes and other factors that can lead to disease or affect a person's response to medications. The project will include background, medical, lifestyle, and genetic information for as many as 500,000 Kaiser Permanente Northern California members who are representative of the population. The data will be combined with their medical history records in a database that will allow research on many hereditary diseases, including cancer, heart disease, Alzheimer's disease, asthma, diabetes, and reproductive problems.[14]

Similar initiatives are under way at other organizations (http://www.aspe.hhs.gov/PHC/rfi/), such as Intermountain Healthcare in Utah and surrounding states (http://intermountainhealthcare.org/xp/public/) and Partners HealthCare in Boston (http://www.partners.org/).

Registries. No registries of patients or families with genetic diseases include sufficient clinical information to assess whether screening criteria for gene-based testing have been appropriately applied.

Other Sources. None was identified.

Linked Databases. None was identified. Patient-level linkage of clinic and laboratory data would be required.

Screening Tests

Questions of Interest. A monitoring system might answer the following questions related to screening tests:

- What gene-based tests are being used and how frequently are they used?

- What health care providers order gene-based tests and where are these providers located?

- Are there changes over time in the types of tests and the frequency at which they are ordered?

- Are laboratories receiving the information needed to conduct and interpret the test from health care providers?

- Do providers have access to evidence-based information about the validity and utility of gene-based tests?

- Do patients have access to gene-based tests? Are there disparities in access to gene-based tests and are there underserved populations?

- Are referrals for gene-based testing appropriate based on indications such as family history and early signs and symptoms of disease?

Existing Information. For this report, we defined screening tests as tests on asymptomatic people to identify people who have an increased risk of developing or having a disease. Screening tests include population-based screening tests, such as prenatal or newborn screening and carrier screening tests.

Newborn screening programs in the United States began in the 1960s with the development of a screening test for phenylketonuria. States routinely test blood spots collected from newborns for up to 30 metabolic and genetic diseases; the four most commonly included are phenylketonuria, congenital hypothyroidism, galactosemia, and sickle cell disease. To encourage uniform and comprehensive newborn screening throughout the United States, HRSA issued a report in 2005 that recommends screening for 29 conditions.[40] Disorders detected through newborn screening are described at NIH's Genetics Home Reference.[12]

Ten percent of 60 primary care physicians interviewed in 1999 in the DC area had ordered a cystic fibrosis carrier test.[16] Among US obstetricians surveyed in 2003, carrier screening varied by the patient's plans for pregnancy.[24] Strom et al. reported on 335,204 postnatal tests for cystic fibrosis conducted by Quest Diagnostics from July 2001 through 2003.[41] They were unable to separate tests for carrier status and diagnostic tests, but the majority was believed to be carrier screening tests. They identified 10,139 carriers, for a frequency of 1:33, and 239 homozygotes. Four homozygotes who were asymptomatic or had only mild symptoms were apparently identified serendipitously.

Quest also conducts a panel of tests for eight diseases that are prevalent among Ashkenazi Jews, including cystic fibrosis. Quest conducted 2,427 of these panel tests before 2004 (exact time period covered is unclear). Of the tested individuals, 14 percent were carriers of one of the tested diseases and 0.8 percent were carriers of two. No information was reported on trends in the use of testing or on the demographics of those tested.[42]

Strom et al. discussed Quest Diagnostic's experience with Fragile X testing from 1992 through 2006.[36] They conducted 1,192,323 postnatal tests during this period. Testing volume increased 30-fold; the sex ratio of persons tested reversed over the 14-year period. During the first 7 years, the male:female ratio was 2:1, but by the end of the reported period the ratio had become 1:1.5. The relative mean age of males (8.1 years) and females (28.3 years) tested suggests that testing for males is primarily diagnostic, whereas testing for females is primarily carrier testing. Among tested females, the prevalence rates were 0.61 percent for a full Fragile X mutation, 1.69 percent for premutations, and 2.16 percent for gray zone mutations.

In August 2004, 78 percent of obstetricians in the Collaborative Ambulatory Research Network routinely discussed aneuploidy screening with all their pregnant patients.[25] Fifty-five percent offered first trimester screening tests: 7 percent offered maternal serum screening (free β-human chorionic gonadotropin [hCG] and pregnancy-associated plasma protein-A); 15 percent

offered nuchal transparency ultrasound; and 33 percent offered combined maternal serum and ultrasound screening. More than 99 percent offered second trimester screening tests: 49 percent offered the quad screen (MSAFP, hCG, unconjugated estriol, inhibin-A); 44 percent offered the triple screen (MSAFP, hCG, unconjugated estriol); 6 percent offered integrated first and second trimester screening; and less than 1 percent offered the double screen (MSAFP and hCG).

Two Australian studies provide examples of the potential for building a monitoring system through linked databases.[43,44] Muller et al. reported on trends in the utilization of different types of Down syndrome screening from 1995 to 2005.[43] They did not find any significant trend over the period in the total proportion of women screened, which ranged from 69 percent to 79 percent. Major changes in the type of screening occurred, however. Second trimester screening declined from 75 percent of confinements in 1995 to 25 percent in 2005, while first trimester screening increased from 0.8 percent of confinements in 2000 to 49 percent of confinements in 2005.

This article also provided data on testing outcomes, at least in regard to clinical validity.[43] Changes in the type of testing did not affect sensitivity; 74 percent of cases were detected overall, with no significant trend over time. By 2005, however, fewer than half as many invasive tests were performed to diagnosis one Down syndrome fetus (57 percent reduction) or one aneuploid fetus (68 percent reduction) than at the peak in 1996.

O'Leary et al. reported on the clinical validity of first trimester screening for Down syndrome.[44] The screening protocol used 1:300 risk of Down syndrome as a cutoff. The detection rate for Down syndrome was 83 percent and the false-negative rate was 1 in 2,277 tests. The positive predictive value was 1 case for every 17.5 positive screening tests for Down syndrome and 1 case for every 3.8 positive screening tests for all congenital defects. It identified 25 percent of all congenital defects.

In the United States, Benn and Ying examined the benefit of adding a screening protocol for Turner Syndrome (45, X) to existing screening protocols for chromosome abnormalities.[45] Detection did not improve, and the false-positive rate was significantly increased.

Data Sources. Data on the utilization of screening tests come from several sources.

Public Health Databases. All states maintain a newborn screening database; some have a genetic services plan.[11] The National Newborn Screening and Genetics Resource Center (NNSGRC), which is a cooperative agreement between HRSA's Maternal and Child Health Bureau (MCHB), Genetic Services Branch, and the University of Texas Health Science Center at San Antonio, Department of Pediatrics, provides data on newborn genetic tests. NNSGRC maintains the National Newborn Screening Information System (NNSIS) database, which is a state self-reported Internet-based, real-time information collection and reporting system for capturing state and territorial newborn screening information. The system provides reports on the numbers of cases and births and puts out a series of reports on newborn screening programs.

Electronic Medical Records. Electronic medical records systems may provide data on screening tests. The Veterans Administration (VA) uses a single electronic medical record, called VistA, in more than 1,300 sites. A similar system is used in Department of Defense facilities.

Linkage between the two systems is currently underway. All VA patient records are accessible to any facility within the system through Web-based access. MEDCIN®, the knowledge engine underlying the VA and Department of Defense systems, has more specific coding of laboratory tests than other coding systems, but the codes may still not be specific enough for monitoring genetic tests. The system incorporates the patient's health history, medications, laboratory testing, and other diagnostic testing. The KP Health Connect™ electronic medical record system was described above under screening criteria.

Clinical Databases. Clinical databases were not used for the studies of screening test utilization that we reviewed. Clinical databases, such as the CRN database described above, could provide information on the course of gene-based screening and testing for patients, as in the two Australian studies discussed above.[43,44]

Laboratory Databases. Four studies were drawn from a testing laboratory database.[36,41,42,45] Laboratory databases contribute key data to monitoring gene-based tests, but they have both strengths and drawbacks. For some gene-based tests, only one laboratory provides the test, so their data on that test are population based. Other tests are offered by many laboratories; thus, including data from all providers of a certain test may be difficult.

Accessing genetic testing data to be combined for surveillance may face additional difficulties. For example, these laboratories tend not to be open to sharing data, partly because of patient privacy issues but also because of their own concerns about data control and proprietary issues (e.g., the possibility that others might be able to compute market share is a barrier to sharing).

Although compiling information about genetic testing data from all laboratories was neither possible nor intended given the scope of this review, briefs about selected laboratories involved in genetic testing are outlined below. The GeneTests website (www.genetests.org) provides probably the most complete list of laboratories conducting gene-based testing. Two large commercial laboratories conducting gene-based tests are LabCorp, including their subsidiary National Genetics Institute, and Quest Diagnostics.

LabCorp is one of the world's largest clinical laboratories and offers a broad range of genomic tests. Their molecular genetics testing center was involved with the original research and subsequent standardization of Her-2Neu testing for breast cancer. The National Genetics Institute (NGI), a wholly owned subsidiary of Laboratory Corporation of America® Holdings (LabCorp®) provides advanced clinical genetics testing services for blood screening, medical testing, and clinical research.[46] Like most other testing laboratories, NGI does not make any data public.

Quest Diagnostics is among the nation's leading providers of gene-based medical testing, with a focus on infectious diseases, oncology, and hereditary conditions. It was the first commercial laboratory to offer several gene-based tests, including HIV resistance testing, national availability of the HER2 gene test for breast cancer patients, and human papillomavirus reflex testing. Quest's key areas of focus include commercialization of diagnostic applications in functional genomics (the analysis of genes and their functions) and proteomics (the discovery of new proteins).[47]

Administrative Databases. CMS allows limited use of beneficiary claims data under a series of procedures, depending on the level of personal identifiable information. The available data do not appear to contain the specific or longitudinal information needed to monitor gene-based testing. Administrative codes, such as International Classification of Diseases (ICD), are usually not specific enough for such purposes.

The 38 Blue Cross/Blue Shield plans each have a database to track claims filed by health care providers for services rendered to plan enrollees. In addition, these plans examine utilization, the appropriateness of care delivery, and outcomes of care. Some plans flag genetic tests using S codes (temporary codes until Common Procedural Terminology [CPT] codes become available) and BRCA diagnoses. Each plan's database includes the standard claims information—laboratory codes, S-Codes (temporary), CPT codes and some aggregators for procedural codes that identify a set of procedures that constitute a specific genetic test. Other modifiers provide more information about the test and the condition for which the test was given, but the plans do not necessarily include them in their information systems. Federal Employee Health Benefit Plan (FEHBP) databases provide information on all federal employees and the covered health services that they receive. Non-FEHBP data could be provided, but access would have to be investigated with each individual plan.

Scheduled Surveys. No article with data on utilization of screening tests used scheduled surveys. Several states have attempted to use PRAMS to examine the use of newborn hearing screening but, to our knowledge, no attempt has been made to validate the results.

The Medicare Health Outcomes Survey is used to gather valid and reliable health status data on Medicare managed care beneficiaries for use in quality improvement activities, plan accountability, public reporting, and improving health. All managed care plans with Medicare Advantage (MA) contracts must participate. A random sample of Medicare beneficiaries, who were continuously enrolled for a 6-month period, is drawn from each participating plan and surveyed every spring (i.e., a survey is administered to a different baseline cohort, or group, each year). Two years later, these same respondents are surveyed again for follow-up measurement. Effective in 2007, the plan sample size was increased to 1,200. This survey could potentially be linked to Medicare claims data to monitor outcomes in patients who have had a gene-based test, although the sample size is probably insufficient to address outcomes of uncommon gene-based disorders. Additionally, young people are poorly represented in this population.

Point-in-Time Surveys. Single point-in-time surveys were the source of most of the data on US screening test utilization.

Ongoing Research Cohorts. None was identified.

Registries. None was identified with data on gene-based testing.

Other Sources. NCI has launched the Cancer Biomedical Informatics Grid™ (caBIG™) initiative to link researchers, physicians, and patients throughout the cancer community. caBIG™ is a voluntary network that enables the collection, analysis, and sharing of data and knowledge along the entire research pathway from laboratory bench to patient bedside. This

network could serve as a source of information on gene-based testing or advise on the collection of information regarding gene-based testing.[48] None was identified.

Linked Databases. Both Australian studies cited earlier linked laboratory and public health databases, including data from antenatal and neonatal screening programs, cytogenetics laboratories, ultrasound facilities, birth and abortion registries, and birth defects registries.[43,44] Although not reported in either analysis, the databases used in these studies should also be able to provide information on pregnancy outcome and, possibly, on complications and treatment decisions and outcomes for live born infants.

Laboratory testing in the United States is less centralized, which would complicate any attempt for a similar data linkage. Newborn genetics screening and birth certificates are linked in some states, although not many.[49] Further linking these data to those of genetics laboratories and ultrasound facilities, particularly in regions that have one or two major providers, might be possible.

Nineteen states have a genetic services plan listed in the NNSIS. Some genetics services plan to discuss integration of data, but the outlook was not promising in the states reviewed (Georgia and Texas).

Interpretation and Followup of Screening Tests

Questions of Interest. A monitoring system might answer the following questions related to the interpretation and followup of screening tests:

- Are referrals for gene-based testing appropriate based on indications such as family history and early signs and symptoms of disease?

- Are patients receiving appropriate counseling about the implications of taking the tests and about test results?

- Are health care providers receiving from laboratories the information needed to interpret correctly test results and implications for their patients?

Existing Information. Physician uncertainty about risk factors for disease and the magnitude of risk associated with susceptibility genes may contribute to inappropriate screening practices. In one study, physicians who treated patients with Alzheimer's disease were surprisingly inaccurate in estimating the lifetime risk of this disorder in patients with or without an affected parent and in identifying risk factors for the disease;[18] 30 percent to 45 percent identified risk factors associated with other forms of dementia as a possible cause of Alzheimer's disease.[18]

Cleary-Goldman et al. reported the following levels of knowledge among surveyed obstetricians:[25] 46 percent knew the risk of a 40 year-old patient having a liveborn child with Down syndrome; 23 percent knew the risk of aneuploidy associated with a first trimester cystic hygroma; 40 percent could identify the second trimester maternal serum results associated with an increased risk of Down syndrome; and 30 percent could identify how much a woman's age-related Down syndrome risk drops if an ultrasound of the fetus is normal.[25]

Lynch et al. reported a case in which multiple family members had colorectal cancer, strongly indicating a diagnosis of HNPCC. [38] The proband was diagnosed with colorectal cancer at age 45. At age 57, she developed vaginal bleeding but was told by her physician, who apparently did not recognize her high risk of endometrial cancer, that the bleeding was the result of a "vaginal yeast infection."

Many physicians are aware of their uncertainty in interpreting genetic screening and testing information for their patients. In one study, only 29 percent of all physicians felt qualified to provide genetic counseling regarding cancer susceptibility testing to their patients.[27] Specialists may be more comfortable with genetic testing; 84 percent of oncologists felt qualified to recommend genetic testing, compared with 58 percent of other tertiary care providers and 41 percent of primary care providers.[27]

Most obstetricians felt somewhat qualified to counsel patients about prenatal genetics. In one study, 89 percent said they were somewhat or well qualified to provide general prenatal genetic counseling, 83 percent felt qualified to provide counseling to patients at elevated risk of fetal aneuploidy, and 80 percent felt qualified to provide counseling to patients screening positive for fetal aneuploidy.[25] Fewer obstetricians were comfortable interpreting genetic tests for cancer; the proportion that were not very confident or not at all confident interpreting tests was 65 percent for breast cancer, 61 percent for ovarian cancer, and 77 percent for colon cancer.[26] Two-thirds reported they would refer a patient with a family history of an untreatable disease to a medical geneticist or a genetic counselor.[26]

Physicians who felt qualified to recommend cancer susceptibility testing were twice as likely to have used testing.[20] Thirty-four percent of primary care physicians in Texas felt that difficulty in interpreting genetic tests was a barrier to their use.[23]

Referral to a genetic counselor or geneticist could compensate for physician uncertainty regarding genetic risk and genetic testing, and many physicians use these services. Twenty-one percent of obstetricians provided counseling themselves, 54 percent referred women to a geneticist, and 22 percent did both.[26] Forty-eight percent of primary care physicians in Cincinnati would refer patients who were candidates for breast cancer susceptibility testing to a genetic counselor and 34 percent would refer them to a geneticist.[21] Twelve percent did not know to whom they would refer such patients.[21] Eighty-five percent of New York gastroenterologists said they would provide genetic counseling before testing for hereditary cancer susceptibility. Twenty-nine percent would provide the counseling themselves, and 51 percent would refer to a genetic counselor. Only 26 percent, however, would have referred a hypothetical HNPCC patient for genetic counseling.[17]

A lack of available genetic counseling creates a substantial barrier to genetic testing. Ninety-one percent of physicians agreed that patients should not have genetic testing unless genetic counseling was available.[27] Physicians who reported that local counseling and testing were available were 50 percent more likely to have used cancer susceptibility testing.[20] Sixty percent of obstetricians have access to community-based counselors, and only 1 percent do not have access to any counselors.[26] However, 44 percent of primary care physicians in Texas reported that genetic evaluation services were not available.[23] Among physicians who requested testing

for or information about breast cancer susceptibility testing, only 12 percent did not have access to genetic counseling services.[19]

Data Sources. Existing information regarding the interpretation and appropriate followup of newborn screening tests is based primarily on survey data, with some information drawn from clinical databases described above.

Public Health Databases. State newborn screening programs may be able to provide information on the interpretation and followup of newborn screening tests.

Electronic Medical Records. Electronic medical records, such as KP HealthConnect[TM] discussed above, may provide additional information on the interpretation of gene-based screening tests, but such information will probably be imbedded in text fields of physicians or other notes in the record. Such information may be difficult to extract.

Clinical Databases. Clinical databases could potentially provide information on the proportion of patients with abnormal screening test results who received referrals for genetic counseling or for follow-up testing. It would be difficult, however, to determine whether patients received appropriate advice on followup and their adherence to recommendations. This level of detail would probably require medical record abstraction, such as was used by Lynch et al. to provide the HNPCC case scenarios.[38]

Laboratory Databases. Laboratory databases provide little information on how test results are used or interpreted. If the same laboratory is used for the screening test and the diagnostic test, it may be possible to infer whether or not the test was interpreted correctly.

Administrative Databases. None was identified, although some information may be available from the Blue Cross/Blue Shield FEHBP database described under screening tests.

Scheduled Surveys. The NCI physician survey and the NCHS NAMCS could be potential sources of data on physician interpretation of gene-based screening test results.

Point-in-Time Surveys. With the exception of the case scenarios by Lynch et al.,[38] all the existing information on physician interpretation and followup of screening tests is based on surveys of physicians.

Ongoing Research Cohorts. See discussion under screening criteria regarding the Center for Human Genetics at the Marshfield Clinics Research Foundation and Kaiser Permanente Research Program on Genes, Environment and Health.

Registries. Monitoring utilization and followup for some gene-based screening tests with registry data that include clinical data may be possible.

The SEER registries for cancers, for example, could provide excellent data on the number and types of tumors (diagnoses) in a selected geographic area. Nine SEER registries (SEER 9) located in Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, and Utah have been collecting data since the 1970's.[50] SEER limited-use data include incidence and population data categorized by age, sex, race (white, black, and

other), year of diagnosis, and geographic areas (including SEER registry and county). The data set includes cases diagnosed from 1973 through 2004 and contains one record for each of 3,405,500 tumors. SEER registries keep individual identifiers, for purposes of doing active followup on survival status. SEER registries, and generally all state and regional cancer registries in the United States, ascertain type of treatment, not just pathology data. The well-developed registries across the country, such as those that participate in SEER, have capabilities to ascertain more detailed clinical data via patient and provider interview, chart reviews, and lab records via special studies on cancer cases.

Other Sources. No published studies were identified.

Linked Databases. Linked clinical and laboratory databases would allow some assessment of how screening test results are interpreted, although differentiating between physician or patient misinterpretation and a patient decision not to follow recommendations may be difficult.

Diagnostic Testing

Questions of Interest. A monitoring system might answer the following questions related to diagnostic testing:

- What tests are being used?

- By whom?

- Where?

- At what frequency?

- Are there changes over time in the number and types of gene-based tests offered and used?

Existing Information. We classified diagnostic tests as those used to identify or confirm a change in gene sequence, either inherited or acquired. (This definition differs from the more common usage of the term "diagnostic test" as a test to confirm a disease or condition, not a genetic sequence.) These tests include susceptibility tests for familial cancers or other adult onset diseases, tests for single gene disorders, and tests to diagnose cancer types or to help determine treatment options.

Seven articles reported on physician use of tests for genetic susceptibility to cancer or referrals for evaluation for such testing.[16,17,19-23] The data reported ranged from 1996 to 2002 and were from varied regions in the United States. Six articles had an unselected sample of physicians. They reported that 20 percent to 50 percent of physicians had ordered a test or referred patients for evaluation (see Table 4).[16,17,20-23] A seventh study drew its sample from physicians who had requested information on or ordered a BRCA1/2 test; 56 percent of these physicians had ordered a BRCA1/2 test.[19]

Table 4. Utilization of cancer susceptibility testing

Cancer Type	Percentages Who Reported Utilization			Location	Year of Data (reference)
	Referred for Testing or Evaluation	Ordered Test	Referred or Ordered Test		
Total	38.8	26.1[*]	37.3[*]		
Colorectal cancer		31[†]	31	New York	1998[17]
Breast cancer		56[‡]	56	Pennsylvania	1996[19]
Breast cancer	51		51	Cincinnati	2002[21]
Breast cancer		35	35	Maryland, Virginia, and District of Columbia	1999[16]
Cancer (unspecified)	39	20	39	Texas	2001[23]
Cancer (unspecified)	26.4	7.9	31.1	US National	2000[20]
Cancer (unspecified)		36.7	36.7	Pennsylvania, New Jersey	2000[22]

[*] Results from Cho et al.[19] was not included in totals because the study was limited to physicians that had requested information or ordered a genetic test.
[†] Based on hypothetical case, not actual practice.
[‡] Sample was biased towards testing.

Tests may not always be ordered appropriately, however. Only 21 percent of New York gastroenterologists who analyzed a case study in HNPCC recognized that an affected relative needed to be tested before testing the unaffected proband would be informative.[17] Thirty-one percent would recommend that the proband receive genetic testing.[17] Fifteen percent of physicians who sought information or ordered genetic testing for breast cancer susceptibility reported ordering a test for breast cancer susceptibility in the absence of family history at a patient's request.[19]

Chase et al. surveyed physicians who had treated Alzheimer's disease patients who were enrolled in a genetic linkage study.[18] Of these, 18 percent had ordered an APOE test and 5 percent had ordered a presenilin 1 test. Physicians were unclear about the appropriate use of these tests, although neurologists and psychiatrists were better informed than other physicians. Fifty-three percent of neurologists and psychiatrists correctly answered a set of three questions designed to assess appropriate use of APOE testing, whereas only 32 percent of other physicians did. Likewise, in the smaller group of physicians who were aware of the presenilin 1 test, 83 percent of neurologists and psychiatrists and 46 percent of other physicians correctly answered questions regarding its appropriate use.[18]

In 2004, 88 percent of obstetricians routinely offered amniocentesis and 44 percent offered chorionic villus sampling (CVS) for pregnancies at elevated risk for any genetic abnormality.[25] Two percent offered diagnostic testing to all their patients. Forty-nine percent of obstetricians did not routinely offer CVS, but only 4 percent of obstetricians did not routinely offer amniocentesis.[25] Kaiser Permanente, a large national HMO, had more than 11,000 cytogenetics tests in 2003 and more than 24,000 molecular laboratory tests.[13]

In Australia, invasive diagnostic testing declined significantly over a decade, from 11.5 percent in 1996 to 7.6 percent in 2005, as first trimester screening tests became more widespread.[43] The decline was seen in women 35 years of age and older and in the general

pregnant population. At the same time, the number of invasive tests required to detect one case of Down syndrome dropped from 1 in 47 in 1996 to 1 in 15 in 2005.[43]

Quest Diagnostics reported on 445 prenatal tests for cystic fibrosis, of which 20 percent were referred because both parents were carriers, 22 percent because ultrasound revealed echogenic bowels, and 58 percent for other or unknown indication.[41] The prevalence of affected fetuses was 1 in 5 when both parents were carriers and 1 in 100 among fetuses with echogenic bowels. None of the 225 prenatal tests that were performed for other reasons identified any cases of cystic fibrosis. In 145 of the tests performed for other reasons (63 percent), an invasive genetic test could have been avoided by carrier testing of the parents or consideration of the utility of the possible tests results.

Between 1992 and 2006, Quest Diagnostics performed 59,707 postnatal Fragile X tests on males, which the authors believed to be primarily diagnostic tests, and 307 prenatal Fragile X screening tests.[36] Among boys tested postnatally, the prevalence of a full Fragile X was 1.4 percent; of a premutation was 0.56 percent; and of gray zone mutations was 0.87 percent. Of the 307 prenatal tests, 165 infants were male and 142 female. Among the male fetuses, 68.5 percent were normal, 12.1 percent had gray zone mutations, 10.3 percent had premutations, and 9.1 percent were affected. Among the female fetuses, 70.4 percent were normal, 12.7 percent had gray zone mutations, 11.3 percent had premutations, and 0.7 percent was affected.

During the late 1990s, Gringras questioned pediatricians in the Thames region in Britain about the diagnostic protocol they would use to diagnose a hypothetical 3-year-old boy with developmental delay of no obvious cause.[51] The two most common tests they would order were karyotyping (86 percent of respondents) and Fragile X testing (81 percent). The other investigative steps were thyroid function (43 percent), metabolic studies (36 percent), and measurement of creatinine phosphokinase (21 percent), mucopolysaccaride (18 percent), urea and electrolytes (18 percent), and calcium and phosphate (18 percent). Other tests were mentioned by less than 10 percent of pediatricians.

Rauch et al. investigated the diagnostic yield of several genetic approaches to investigating developmental delay.[52] Targeted analyses based on dysmorphic findings yielded a diagnosis in 20 percent of cases; standard karyotyping identified the diagnosis in about 16 percent of cases.

Sequist et al. reported on their experience with epidermal growth factor receptor (EGFR) tumor mutation testing for patients with non-small-cell lung cancer. Of 278 patients, 68 (24 percent) were EGFR positive.[53] The prevalence of EGFR-positive patients was higher than in previous reports. Of the 68 EGFR+ patients, 61 patients had one mutation, 6 patients had two mutations, and 1 patient had three mutations. Thirty-nine percent of patients had an in-frame deletion in exon 19, and 33 percent had point mutations in exon 21. Four patients had mutations (T790M and L858R) that have been associated with resistance to tyrosine kinase inhibitors. Although T790M is thought to be acquired in response to treatment with tyrosine kinase inhibitors, only one of the patients with the mutation had been exposed to such treatment.

Data Sources. *Public Health Databases.* Existing information on the utilization of newborn screening diagnostic testing was drawn from laboratory and clinical databases, ongoing research

cohorts, and point-in-time surveys. See discussion under screening tests of the National Newborn Screening Information System (NNSIS) database.

Electronic Medical Records. The utilization data for Kaiser were abstracted from their KP HealthConnectTM™ electronic medical record discussed above.[13] When it is fully implemented, this database could provide additional information on diagnostic testing as well.

Clinical Databases. Extracting the number of tests ordered and the demographic characteristics of the patients from clinical databases may be possible. As we discussed under interpretation and followup of screening tests, clinical database abstraction, such as was used to identify the case studies presented by Lynch et al.[38] would be needed for more detailed information.

Laboratory Databases. Laboratory databases were used for some of the articles reported in this section.[36,41,43,44,53] The strengths and weakness of laboratory databases for monitoring gene-based diagnostic testing are the same as those for screening tests. The demographic and clinical information available is very limited. Linking test results for a given patient is typically impossible because screening and diagnostic tests may go to different laboratories. Clinicians may also repeat abnormal tests at a second laboratory, so the same patient may be counted twice.

As discussed under screening tests, gaining the cooperation of laboratories for a monitoring system for gene-based tests may be difficult. For example, Myriad Genetics, a large gene-based testing laboratory in Utah, does not provide data for public or research use (personal communication, Eric Rosenthal, Myriad Genetics, September 4, 2007).

Administrative Databases. See discussion of Medicaid and Medicare data above. Information on diagnostic tests would be available from the Blue Cross/Blue Shield databases as described under screening tests.

Scheduled Surveys. The NCI physician survey and NAMCS are potential sources of data on physician interpretation of gene-based screening test results. The Medicare Health Outcomes Survey described above could potentially be linked to Medicare claims data to monitor outcomes in patients who have had a gene-based test, but again the sample size may not be sufficient to allow analysis of uncommon disorders.

Point-in-Time Surveys. Much of the existing information on diagnostic gene-based testing was derived from point-in-time surveys.

Ongoing Research Cohorts. Kolibianakis et al. presented data on the utilization and outcomes of amniocentesis and CVS from an ongoing cohort study of outcomes after assisted reproductive technologies.

Registries. Registries of patients with specific disorders could include data useful for monitoring gene-based testing, but as noted earlier, they are mainly research tools and not aimed at public health monitoring. The NCI Breast and Colon Cancer Family Registries, for example, include information on family history, epidemiologic, clinical data, genotypes and updates on cancer recurrence, morbidity, and mortality in participating families. These registries facilitate and support interdisciplinary and population-based research on the identification and

characterization of breast, ovarian, and colorectal cancer susceptibility genes, with particular emphasis on gene-gene and gene-environment interaction research.[54] Data dictionaries were not available online. SEER limited-use data were described earlier.

Other Sources. Some companies track use of their tests. For example, Genomic Health is collaborating with United Healthcare to track use of its Oncotype DX test for women with breast cancer (personal communication, Dr. David Veenstra, Institute for Public Health Genetics at the University of Washington, July 26, 2007). Companies that provide services to the pharmaceutical industry may also maintain databases of test utilization. For instance, Quintiles Informatics sells a large, real-time, patient-level claims database of more than 3 billion pharmacy and medical transactions representing health care experiences for more than 150 million de-identified patients in the United States. Such information is usually quite expensive and its use is highly restrained, however (personal communication, Dr. Claude Hughes, RTI International, July 18, 2007).

The College of American Pathologists (CAP) through its CAP Laboratory Accreditation Program, maintains a large database that might be a significant source of data on genetic testing. These CAP data reflect an annual survey of nearly 10,000 laboratories, including those that conduct acquired disease and inheritable disease genetic testing. The CAP Accreditation and Lab Improvement data could be used to monitor the volume of the genetic tests.

Linked Databases. The report by Muller et al. on prenatal testing for cytogenetic abnormalities used the linked clinical, laboratory, and public health databases discussed under screening tests. As noted earlier, for most types of gene-based tests, such a linked database would be very difficult to construct in the United States.

Interpretation of Diagnostic Tests

Questions of Interest. A monitoring system might answer the following questions related to the interpretation of diagnostic testing:

- Do providers have access to evidence-based information about the validity and utility of gene-based tests?

- Do patients have access to gene-based tests?

- Are there disparities in access to gene-based tests and are there underserved populations?

- Are health care providers receiving from laboratories the information needed to interpret correctly test results and implications for their patients?

- Are patients receiving appropriate counseling about the implications of taking the tests and about test results?

Existing Information. We found little information on the interpretation of diagnostic tests other than what can be inferred from the barriers that physicians see to testing and their recommendations for patient management and family education. We discuss barriers to testing

related to interpretation in this section. Patient management and family education are discussed below.

In one study, only 16 percent of primary care providers knew that more than 50 percent of patients with the HNPCC gene will develop colorectal cancer; 40 percent thought the risk was less than 50 percent, and 44 percent did not know.[22] The actual risk is estimated to be 72 percent to 80 percent.[22]

Of a national sample of 1,215 physicians, 45 percent agreed with the statement that the risk of cancer in patients with a positive genetic cancer susceptibility test is not clear, and 75 percent agreed that clear guidelines for managing someone with a positive test result were not available.[27] Among a sample of primary care physicians in Texas, 60 percent identified lack of guidelines for patient management as a barrier to greater use of cancer susceptibility testing, and 36 percent considered difficulty in interpretation as a barrier.[23] In a study of obstetricians, 36 percent were unclear about how the results of cystic fibrosis or breast cancer susceptibility tests would change their management of the patient.[16]

A substantial minority of physicians also have concerns about the clinical validity of these tests. In one study, 25 percent of physicians said that genetic tests for cancer susceptibility had too many false positives, false negatives, or ambiguous results.[27] Primary care providers in Texas were also concerned about false positive (44 percent) and false negative (30 percent) results. Twenty-six percent of obstetricians were uncertain of the sensitivity and specificity of cystic fibrosis or breast cancer susceptibility tests.

Data Sources. *Public Health Databases.* None was identified.

Electronic Medical Records. Inferring some information on interpretation from electronic medical records may be possible.

Clinical Databases. None was identified.

Laboratory Databases. None was identified.

Administrative Databases. None was identified.

Scheduled Surveys. None was identified.

Point-in-Time Surveys. Current data on the interpretation of diagnostic tests were all drawn from physician surveys. Such surveys are likely to remain the primary source of such data.

Ongoing Research Cohorts. None was identified.

Registries. None was identified.

Other Sources. None was identified.

Linked Databases. None was identified.

Impact on Treatment Choices and Short-Term Clinical Outcomes

Appropriate Clinical Management of Patients

Questions of Interest. Some questions of interest that might be answered by a monitoring system related to the clinical management of patients are:

- Do patients have access to followup and interventions that might be recommended based on test results?

- Are patients offered options for different treatments or preventive measures?

Existing Information. Physician knowledge and practice of appropriate clinical management for patients with familial cancer syndromes are uneven. Only 16 percent of gastroenterologists in New York would have recommended colonoscopy every 1 to 2 years for a hypothetical HNPCC patient.[17] Lynch et al. reported a case in which a member of an HNPCC family repeatedly informed her physician that she had a hereditary cancer risk but was told each time that she was fine and should not worry. Repeated letters from the authors were required before she was scheduled for a colonoscopy, at which time she was diagnosed with a Stage II adenocarcinoma.[38] Only five of nine North American registries of families with familial adenomatous polyposis recommended that gene-negative family members be discharged; the other four registries recommended a cancer screening schedule that was less frequent than for mutation carriers but more frequent than for the general population.[55]

The European Familial Breast Cancer Collaborative found that increased cancer screening was moderately effective in detecting tumors in women at risk for familial breast cancer who complied with the recommended screening regime. Seventy-five percent of tumors were detected by examination and 57 percent were detected by mammography.[52] For women under age 59 at diagnosis, 68 percent of their tumors were detected by examination and 45 percent by mammography.

Tinley et al. examined the compliance of family members who had a BRCA1/2 mutation or were at 50 percent risk of having a mutation with the cancer screening recommendations provided by cancer geneticists during family information sessions.[56] These recommendations include monthly breast self-examination, semiannual clinical breast examination, annual mammograms for breast cancer screening, and annual ultrasound and semiannual CA125 screening for ovarian cancer. The personal physicians of 89 percent of patients with a BRCA1/2 mutation also recommended annual mammography. Only 22 percent of patients with a BRCA1/2 mutation and a family history of ovarian cancer had a personal physician who recommended regular ultrasound and CA125 screening. Physician recommendation was highly associated with patient adherence to the screening protocol. Among patients whose personal physician also recommended annual mammography, 79 percent adhered to recommendations; by contrast, among women whose personal physician did not make the recommendation, only 10 percent complied.[56] Similarly, 66 percent of patients whose physician recommended ovarian cancer

screening followed the recommendation, but only 6 percent of those whose physician did not recommend the screening received the appropriate screening.

Data Sources. *Public Health Databases.* None was identified.

Electronic Medical Records. See discussion of KP HealthConnect™ under Screening Tests.

Clinical Databases. All the data in this section were based on clinical databases at one or more treatment institutions.

Administrative Databases. None was identified.

Scheduled Surveys. None was identified.

Point-in-Time Surveys. None was identified.

Ongoing Research Cohorts. None was identified.

Registries. None was identified.

Other Sources. None was identified. The NIH Pharmacogenetics Research Network (PGRN) conducts and collates basic and clinical research on pharmacogenetics, but the data are not geared toward public health surveillance or health outcomes research. The PGRN comprises 12 independently funded interactive research groups, each having a focus in an identified area of pharmacogenetics. The PGRN is accomplishing its mission by conducting studies of variation in human genes relevant to pharmacokinetics (drug disposition) and pharmacodynamics (drug action) and the relationship of such variation to drug response phenotypes. The resulting data are deposited into the PharmGKB knowledge base, which contains data and information accumulated in the field and contributed by researchers both within and beyond the network.

Linked Databases. Laboratory data linked with clinical records or electronic medical records could potentially provide information on whether patients are receiving appropriate management. As noted before, however, electronic medical records are in their infancy, and linking databases at the patient level would be extremely difficult.

Appropriate Family Followup

Questions of Interest. A monitoring system might answer the following question related to followup with the families of patients:

- Do families have access to followup and interventions that might be recommended based on test results?

- Are family members offered options for different treatments or preventive measures?

Existing Information. A recent review found that family members of patients with HNPCC mutations were more likely to undergo colorectal cancer screening and had improved outcomes due to the screening, regardless of whether they had an HNPCC mutation or not.

Data Sources. *Public Health Databases.* None was identified.

Electronic Medical Records. None was identified.

Clinical Databases. Clinical database from specialized genetics clinics or genetic counseling clinics may provide information on family followup.

Laboratory Databases. None was identified.

Administrative Databases. None was identified.

Scheduled Surveys. None was identified.

Point-in-Time Surveys. None was identified, but point-in-time surveys of probands or family members could potentially provide information on family followup after gene-based testing.

Ongoing Research Cohorts. None was identified. Ongoing research cohorts may provide information on family followup, but the findings may not reflect routine clinical practice.

Registries. None was identified, although registries of families with the condition of interest could potentially provide information on whether the families are receiving appropriate followup, either through data collection by the registry or by providing a sampling frame for surveys.

Other Sources. None was identified.

Linked Databases. None was identified.

Health and Long-Term Outcomes

Questions of Interest. A monitoring system might answer the following question related to health outcomes of patients and their families:

- What is the impact of gene-based tests on patients, families, and the health care system?

Existing Information. Kolibianakis et al. reported that, among women who became pregnant using intracytoplasmic sperm injection, 3.7 percent suffered a fetal loss and 0.9 percent experienced a loss with amniocentesis.[57] These investigators did not include a control group, however, or otherwise compensate for the increased likelihood of first trimester fetal loss unrelated to treatment.

Evans et al. compared the survival of patients seen in their family history clinic to that of patients seen in their surgery clinic, which were considered sporadic cases.[58] The hazard ratio of death from breast cancer for family history clinic patients compared with surgery clinic patients was 0.24 (95% confidence interval [CI]: 0.09 – 0.66), and the hazard ratio for recurrence of breast cancer was 0.25 (95% CI: 0.11 – 0.57).[58] The authors concluded that the intensive screening for patients identified at high risk both delayed death and increased disease-free survival. Van Roosmalen et al. found that BRCA1/2 testing and disclosure of a positive test

result had an adverse effect on the well-being of women whether or not they had cancer at the time.[59]

A recent review identified three studies[60-62] that addressed the potential psychological effects associated with HNPCC testing among patients and family members.[39] Gene-based testing did not affect depression in either patients or family members. One study found that anxiety and cancer-related distress decreased after testing,[62] while another found no effect.[60] One study found that 33 percent of the six mutation carriers in the study experienced extreme guilt that they may have passed the mutation onto their children.[60] Limited information is available on the impact of HNPCC testing on prognosis. One study suggested that prognosis was better among mutation carriers than nonmutation carriers with microsatellite stable tumors,[63] and another found increased survival among colorectal cancer patients who met the Amsterdam 1 criteria for familial cancer compared to those who met the Japanese 1 criteria for sporadic cancer.[64]

Among 92 patients actively treated for non-small-cell lung cancer after having EGFR mutation testing, 15 of 28 EGFR mutation carriers responded to tyrosine kinase inhibitors, whereas none of the 31 EGFR mutation-negative patients treated with these inhibitors responded.[53] The EGFR-positive and EGFR-negative patients did not differ significantly in response to chemotherapy. A recent evidence-based review of three gene expression assays for predicting recurrence risk or response to chemotherapy among breast cancer patients, Onco*type* DX™, MammaPrint® and the Breast Cancer Profiling (BCP or H/I ratio) test, found evidence that Onco*type* DX™ provided clinically useful information in predicting which patients would benefit from chemotherapy. The information regarding MammaPrint® or the BCP test was less well developed and no evidence of clinical utility was found for these tests.[65]

In the Netherlands, screening for familial hypercholesterolemia identified 896 new cases in approximately 4 years. After 2 years, 80 percent of patients were receiving treatment, but only 66 percent of those patients achieved target cholesterol levels.[66] Eighty-five percent of the people screened had a positive opinion of the screening program.

Data Sources. *Public Health Databases.* None was identified. Surveillance systems can provide information on health outcomes associated with some gene-based tests, such as prenatal screening, carrier testing, and newborn screening.

Electronic Medical Records. When fully implemented, electronic medical records should be able to provide the best source of data on health outcomes.

Clinical Databases. Most of the data in this section were based on clinical databases at one or more treatment institutions. The Cancer Research Network (CRN) could also provide data on health outcomes associated with gene-based testing. Building on the CRN infrastructure to monitor tests for conditions other than cancer may be possible.

Laboratory Databases. None was identified.

Administrative Databases. Linking records from administrative databases, such as the Medicare and Blue Cross/Blue Shield databases, may make it possible to obtain some information on health outcomes.

Scheduled Surveys. National surveys could provide information on health outcomes associated with gene-based testing only if the test was indicated for a large proportion of the population and the outcome was relatively common.

Point-in-Time Surveys. Point‑in‑time surveys of selected populations could also provide information on a variety of health outcomes following gene-based testing.

Ongoing Research Cohorts. Van Roosmalen used an ongoing longitudinal research cohort to assess the effect of BRCA1/2 testing on women's well-being. Clinical trials of specific diseases may also be a source of health outcomes data.

Registries. Registries, such as the Breast and Colon Cancer Family registries, could provide some information on health outcomes for patients with specific diseases.

Other Sources. The Cancer Genetics Network (CGN) supported by NCI is a national network of centers specializing in the study of inherited predisposition to cancer. The resource is available to the research community at large to support studies on the genetic basis of human cancer susceptibility; integration of this information into medical practice; and behavioral, ethical, and public health issues associated with human genetics. The growing database has information on 24,000 individuals (16,000 families) with cancer and/or a family history of cancer. Data available to researchers include demographic information, relevant medical history, and a four-generation cancer family history on each enrollee. The population enrolled makes possible research on both common and uncommon tumors. The CGN welcomes opportunities to collaborate with research groups; it can also provide data and biospecimens to support independent studies, such as outcomes studies of cancer genetic testing.

Linked Databases. Laboratory records linked with clinical records or with electronic medical records could potentially provide information on whether patients are receiving appropriate management. Although some integrated health care delivery systems have implemented comprehensive electronic medical records that allow linkage of clinical data at the individual level, the further ability to link with specific genetic laboratory test results is just being developed through projects such as the Kaiser Program on Genes, Environment and Health and similar efforts.

Assessment of Current Databases

We found little or no information available on the issues that motivated this project in the first place—i.e., patterns of use of these tests, patients' and clinicians' access to them, and their impact on treatments and health outcomes. For example, no information is available on the utilization of gene-based tests or on trends over time (beyond the obvious, namely, that use is rising simply by virtue of the wider availability of more tests and growing awareness about them among patients and clinicians).

Some research and surveys suggest that knowledge on the part of some providers about the availability and utility of tests may be reasonably widespread and accurate. Examples include obstetricians for prenatal care, gynecologists for breast cancer, and, to some extent, clinicians treating adults or the elderly about Alzheimer's disease tests. Little or nothing is known about the

extent to which patients and their families are aware of tests and knowledgeable about their benefits and harms.

Finally, little is known or available about the impact that either screening or diagnostic tests have on decision making about clinical management (i.e., the processes of care) or the short- or long-term outcomes of such testing, suggesting that at the moment, little can be said about the quality of care involving gene-based testing.

We identified no database that could provide all the information desired about gene-based testing broadly. Linking databases or collating information from multiple sources may be possible to acquire some of the desired information for a given test. For example, newborn screening databases have already been linked to birth certificates in some states. A great deal of information could be obtained if these databases were further linked to information from the public health genetics programs that coordinate followup and to state birth defects and developmental disabilities surveillance programs.

Thus, current databases might be used alone or in combination to address specific questions about gene-based testing. Whether they would provide an adequate "monitoring" capability remains unclear. As discussed in the next chapter, the current limitations to databases are appreciable and would have to be addressed before this supposition could become a reality. This would include implementing a research agenda that could move development and implementation of viable databases forward.

Chapter 3. Recommendations and Future Directions

In this chapter, we first discuss some of the limitations of current databases and issues that affect their utility for the monitoring (or surveillance) purposes that were the chief focus of this project. We also examine the potential influence of current and future developments in health information technologies (IT) and standards development. Subsequently, we discuss the optimal features that will be required to develop and operate database systems for various monitoring and research purposes. Throughout, we have relied not only on our analyses from the data collection activities outlined in Chapter 2 but also, importantly, on the discussions and recommendations stemming from our invitational workshop and on ideas provided by our external peer reviewers.

Given our conclusions documented in Chapter 2 about the low likelihood that *current* databases could be used for the uses examined in this report, we turn to possible research applications that current (or future) databases might support. In the absence of comprehensive, national monitoring of the use and outcomes of genetic tests or other interventions, high-quality research efforts offer an appealing alternative. Thus, in the last part of this chapter, we propose recommendations for high-priority research topics to foster the development of an enhanced infrastructure for monitoring the utilization and outcomes of gene-based tests and related interventions.

Addressing Limitations of Current Databases

Several limitations and characteristics of current databases pose challenges to their use for the purposes of creating any kind of infrastructure to monitor use and outcomes of gene-based applications. We discuss the key problems below. Table 5 provides a summary of the identified problems and potential solutions.

Table 5. Problems and limitations of current databases and potential solutions

Issue	Problems or limitations	Potential solutions
Purposes and structures of databases	Data not collected or structured for surveillance or research purposes	Adoption and enhancement of electronic health records systems
	Difficulties linking individual-level data across databases	Improvements in informatics and health information technology (HIT)
	Databases and gene-based tests are continuously changing as technology improves	Maintain flexibility in HIT standards and methodologies
Population coverage and data integration	Available databases cover only segments of the population or specific conditions	Population surveys
	Longitudinal followup difficult (e.g., individuals switch health plans)	Although not entirely population-based, longitudinal data may be available from some registries and integrated health care databases

Table 5. Problems and limitations of current databases and potential solutions (continued)

Issue	Problems or limitations	Potential solutions
Population coverage and data integration (continued)	Individuals not covered by health insurance may be missed	Obtain data from commercial laboratories
	May miss direct-to-consumer tests	Population surveys
Barriers to data access and sharing	Various approvals may be needed to access and share data (e.g., IRB, HIPAA, data use agreements)	Guidance on requirements of applicable laws and regulations (and harmonization across different organizations/entities)
	Linking data across health plans or organizations may not be possible	Improvements in informatics and health IT to improve linkage
	Organizational resistance to sharing	Provide incentives for sharing data
Coding and interoperability of systems	Lack specific codes for genetic tests and test results	Promote standard coding conventions
	Diverse database architectures inhibit interoperability	National initiatives to promote interoperable information systems (e.g., NHIN, PHIN, AHIC)
Privacy concerns	Individuals reluctant to share personal health information	Garner public support and trust
		Public and provider education
	Laws and regulations to protect privacy and confidentiality	Adopt stringent security, privacy, and confidentiality standards
		Clarify requirements of various laws and regulations (e.g., HIPAA, CLIA)
Resources	Costs to collect data, enhance databases, or create new systems or surveys could be considerable	Weigh benefits versus costs
		Improve collection of data to evaluate clinical utility of tests or interventions
		Consider "phased reimbursement" approaches to cover clinical utility analyses
Information needs of different audiences	Needs differ for the public, health care professionals, payers, and government	The structure, content, and operation of databases will be the same to address the interests of the different audiences, but the desired information and how it is presented will vary by audience

AHIC, American Health Information Community; CLIA, Clinical Laboratory Improvement Amendments of 1988; HIPAA, Health Insurance Portability and Accountability Act of 1996; IRB, Internal Review Board; IT, information technology; NHIN, National Health Information Network; PHIN, Public Health Information Network.

Purposes and Structures of Databases

The majority of identified databases had a primary purpose other than surveillance of or research on genetic testing or gene-based interventions. Predominant purposes were payments and claims processing, service provision, quality improvement, and accreditation. Thus, the data may not be as complete or free of error as might be desired for either of these purposes.

Considerable effort may be required to simply format and edit data to the requirements of a surveillance or research database, not to mention the difficulties in linking individual-level data from different databases. Table 6 provides an example of the complexities involved in identifying patients' test results for one gene-based test (OncotypeDX™) at Kaiser Permanente Northwest.

Table 6. Identifying patient-specific test results for OncotypeDX™ in Kaiser Permanente Northwest*

Background: Genomic Health (Redwood City, CA) has a monopoly on OncotypeDX testing for women with estrogen receptor positive breast cancer, the results of which are used to help women choose whether or not to have adjuvant chemotherapy for Stage 1 tumors.

Current process for identifying test results:

1. Medical oncologists order this test by sending a paper form to the KPNW pathology department.

2. The pathologists prepare the tumor tissue specimens and send them via overnight courier to OncotypeDX.

3. Genomic Health performs the genetic tests and then makes the results available via a secure password-protected website to the ordering oncologist and follows up with a hard-copy report to the ordering physician by FedEx.

4. In order for KPNW research scientists to find these tests, they have to access the KPNW pathology business support databases to trace the payments from KPNW to Genomic Health and also access the pathology department's log of all the tissue specimens sent out for analysis (the department keeps a complete inventory of all tissue specimens received, stored, and sent out by tumor/tissue sample ID number).

5. The tumor registry is used to create the link from the tumor ID number to the patient health record number.

Proposed solution: This is a complex process, fraught with possibilities for error at many steps. KPNW is working on a potential solution to use the HealthConnect electronic medical record to permit the oncologists to order OncotypeDX testing electronically. This would give researchers access to a record in the electronic medical record data warehouse of the physician's order for OncotypeDX testing for each individually identified patient. However, Genomic Health has been reluctant to transmit their test results via a secure data portal into the HealthConnect databases; thus, the scanned documents have to be read and the test results re-entered for each patient. It is not certain if Genomic Health would change their business practices, but a direct secure encrypted data transfer to the Kaiser Permanente national data warehouse would provide a much higher level of privacy protection overall than the current indirect way of accessing patient-specific test results.

ID, identification; KPNW, Kaiser Permanente Northwest.
*Example provided by Mark C. Hornbrook, Ph.D., Kaiser Permanente Northwest, Portland, Oregon

Another consideration is that health databases are dynamic and, thus, any monitoring system must be able to accommodate rapidly changing health information technologies. The general movement toward electronic health records could help foster capabilities for monitoring gene-based applications.[67] Developments that would give individuals greater control over their

personal electronic health records, however, may hamper access to electronic databases for public health surveillance or health services research purposes. Microsoft and other companies are betting that the future will converge on web-based medical records that are completely portable.[68] This scenario raises the issue of who owns the medical record. For example, Microsoft might own the right to sell an individual's key strokes and this may conflict with patient confidentiality rights. Moreover, if patients are placed entirely in control, they might be able to edit their medical records; some may not want to share personal health data or may selectively restrict availability of certain information (e.g., sexual health, psychiatric history and medications).

Not only are databases changing, but the nature and type of genetic tests and gene-based applications are also rapidly evolving. Genome-wide assays that can measure hundreds of thousands of single nucleotide polymorphisms (SNPs) are now a reality and are becoming increasingly affordable. Dealing with the massive volume of data provided by such assays will require significant informatics developments. For example, it is not unrealistic to imagine that electronic medical records or other health databases will incorporate data on an individual's entire genome at various points in their lifetimes. Of course, large genome-wide scans are already being marketed directly to consumers, and these data may never make it into a database other than that of the testing laboratory. Unless arrangements can be established with testing laboratories, information on utilization of direct-to-consumer testing will have to rely on population surveys.

Population Coverage and Data Integration

Data from even large providers and insurers (e.g., Kaiser Permanente, Blue Cross/Blue Shield) cover certain regions and segments of the populations in those regions. Consequently, even if selected population-based analyses were possible in some geographic areas, national analyses are very unlikely to be possible. Moreover, individuals may lose coverage or switch health care plans, making longitudinal followup challenging. Long-term clinical outcomes data are generally easier to derive for patients who do not move to another system. Patients with chronic diseases, such as cancer, tend to keep their coverage rather than move, making monitoring of related gene-based applications more feasible for such conditions than for acute conditions.

In addition, certain systems may be designed to monitor or evaluate genetic testing in a focused area. For example, the Surveillance Epidemiology and End Results (SEER) program at the National Cancer Institute (NCI) covers cancer, but no similar *surveillance* program exists for other diseases. Moreover, in the example of cancer genetic testing, most of the studies are highly specific to a particular cancer. The merged SEER-Medicare database is the most comprehensive database for "regional tracking"—but this is not a national database (because of the noncomprehensive coverage of SEER and restriction of Medicare largely to people 65 years and older).

Many databases are based on reimbursement claims for health care services. Individuals who lack insurance or choose to self-pay (e.g., for confidentiality reasons) would be missed in these administrative databases. Population-based surveillance intended to include these individuals

would probably have to rely on survey data in which respondents are selected from the entire population.

Barriers to Data Access and Sharing

Except for the clinical laboratories, most databases can be shared with those having a legitimate research use. Most of the databases described, however, require individual approvals for projects from an appropriate Institutional Review Board (IRB). They also are likely to employ stringent data use agreements.

Most databases have a set of common data elements that can be used to link to other datasets, at least within their own systems. This may be especially true with respect to data from laboratories, provider networks, integrated delivery systems, and managed care organizations (MCOs). In principle, linking these databases is possible; however, to do this one would need to get approval from individual owners through IRBs or other procedures in place (e.g., Health Insurance Portability and Accountability Act [HIPAA] requirements). Many genetic conditions are rare and raise particular concerns in developing appropriate databases that adequately protect individual privacy and confidentiality.

Coding and Interoperability of Systems

Tracking genetic testing in the US health care system is not possible. Most large data sources, including administrative and provider databases, use Current Procedural Terminology (CPT) codes, which are not specific for genetic tests. Expanding the codes at this point is not possible, other than by using modifiers. As illustrated in Table 7, some health plans flag genetic tests using S codes (temporary codes secured by manufacturers until CPT codes become available). Also, the coding—both CPT and those of the International Classification of Diseases (ninth version, clinical modification; ICD-9-CM)—do not indicate why the test was ordered. These databases also do not include test results.

Table 7. Overview of steps in assigning codes to a new genetic technology*

1. In research studies, the procedure or test is assigned a special "research" procedure code according to the study protocol on who is going to pay for the experimental test.

2. Then, to move out of "experimental" status, a test must be approved by the major payers as being a covered service. At this point, a test may be given an "S" code or a "Miscellaneous" code.

3. Eventually, as utilization rises, it becomes imperative to assign a new unique code to the test by the coding organizations and the test becomes much more visible to researchers who use electronic data warehouses.

*Provided by Mark C. Hornbrook, Ph.D., Kaiser Permanente Northwest, Portland, Oregon

Any national surveillance system will require large-scale participation from providers, MCOs, and insurance plans. Resistance to sharing data may be strong on the part of these organizations, particularly for genetic tests and test results because of privacy concerns for individuals. Data from commercial testing laboratories could be useful but these sources pose substantial barriers. For example, most of the laboratories conducting tests are private (and may

be for-profit entities). In addition, the country has many testing laboratories, and many of them are not exclusive providers of a given test.

Although current claims-based data systems rely primarily on ICD-9-CM and CPT codes, other coding systems may provide greater clinical detail and may be more suitable for exchanging health care information between systems. Efforts are under way to promote standard coding conventions in US health databases to improve data systems interoperability and efficient data exchange.

One such initiative is the Healthcare Information Technology Standards Panel, whose mission is to serve as a cooperative partnership between the public and private sectors for the purpose of achieving a widely accepted and useful set of standards. It seeks to enable and support widespread interoperability among health care software applications capable of interacting in a local, regional, and national health information network for the United States. Comprising a wide range of stakeholders, the Panel will assist in the development of the US Nationwide Health Information Network (NHIN) by addressing issues such as privacy and security within a shared health care information system. Funding for the Panel is being provided by the US Department of Health and Human Services.[69]

The Public Health Information Network (PHIN) is the architecture from the Centers for Disease Control and Prevention (CDC) for advancing fully capable and interoperable information systems in the many organizations that participate in public health and ensuring that these systems connect to broader national health IT activities.[70] PHIN is a national, multi-organizational business and technical architecture for public health information systems. At the core of PHIN are accepted health data and technical standards including Systematized Nomenclature of Medicine (SNOMED), Health Level 7 (HL7), and Logical Observation Identifier Names and Codes (LOINC):

- SNOMED ensures that health care professionals can communicate effectively with each other and both veterinary and human medicine can use a common language

- HL7 defines the structure of the electronic messages sent between health care computer systems.

- LOINC provides a set of universal names and identification (ID) codes for identifying laboratory and clinical test results; it enables receipt of clinical and laboratory information from a variety of sources.

Specifically related to gene-based data, the American Health Information Community (AHIC) formed the Personalized Healthcare Workgroup (PHC WG) with the charge of determining how health information technology can be used for the development of standards for interoperable integration of genomic test information into personal electronic health records. AHIC is a federal advisory body (http://www.hhs.gov/healthit/ahic/healthcare/), chartered in 2005 (and currently being transitioned to a private-public partnership), to make recommendations to the Secretary of the US Department of Health and Human Services on how to accelerate the development and adoption of health information technology.

Privacy Concerns

Protecting the privacy of individuals and the confidentiality of their health information is a major concern; inadequately meeting this need could hamper making record-level linkages across databases. For example, the Food and Drug Administration (FDA) would like to be able to link its currently available databases in a more useful context. Barriers to accessing some of the FDA databases are privacy and proprietary concerns of the companies (e.g., pharmaceutical manufacturers).

Public perceptions may also be a problem. Some individuals may not want to share their personal health information at all, and public support will be needed to address any potential legal issues.

Many government laws and regulations are in place to protect various aspects of individuals' health data privacy, confidentiality, and security. The public and health care providers may be confused about these various laws and regulations and how they interrelate. Substantial clarification is needed on the variations in state laws and regulations regarding transmission of genetic test data such as newborn screening results. This would include clarifying the role of HIPAA, the Individuals with Disabilities Education Act (IDEA), the Family Educational Rights and Privacy Act (FERPA), and the Clinical Laboratory Improvement Amendments (CLIA) in governing appropriate sharing of newborn screening results and other genetic data. Concerns about privacy must also be balanced against considerations of promoting medical progress and equity. If restrictions on data access are too stringent, the data may not be available for conducting research on the utility of genetic applications or for monitoring their utilization.

Electronic databases containing gene-based information on individuals would face many of the same considerations with which the NHIN is now dealing; the National Committee on Vital and Health Statistics has also been addressing some of these issues.[71] In developing any electronic health databases, support from the public will be crucial; this in turn requires the trust and confidence of the public that their health information is protected. Thus, any databases containing personal genetic or health information must have the utmost security and incorporate privacy and confidentiality protections as central features. Public trust will require public and provider education not only of the various regulatory and technological protections built into the databases, but also, and perhaps more importantly, an understanding of the value of the databases in improving their personal health care.

Resources

In developing and maintaining databases, many costs could be incurred at various levels. Among these will be laboratory costs to save data in useable databases, provider costs to implement or convert data systems or to survey patients, and costs to patients or the public for completing surveys or providing additional information. Considerable resources will be required to take existing sources of data and convert them into databases that could be used for public health surveillance or health services research. Converting such data, especially if created originally for other purposes, into standard consistent formats is very time and resource intensive. This problem is magnified to the extent that multiple data sources from different organizations are to be combined. The fixed costs of making the initial conversion may be

substantial; however, the marginal costs of maintaining the databases would be more affordable. Thus, the benefits of certain database improvements may outweigh the costs, particularly for large health plans and MCOs.

Evidence of clinical utility is becoming increasingly important in the decisionmaking of health care payers. The need for evidence to support coverage and reimbursement decisions could potentially be leveraged to support database development and health services research. For example, the Secretary's Advisory Committee on Genetics, Health, and Society (SACGHS) has recommended that public and private health care payers develop "coverage with evidence" or "phased reimbursement" approaches to help cover the collection of clinical utility evidence.

Information Needs of Different Audiences

Development of any database or system of databases for either monitoring or research will require support at multiple levels. Crucial questions are what information is needed, and for whom, for public health surveillance or research applications. In other words, who is the audience? In fact, multiple audiences will be important to reach in these endeavors:

- Public (taxpayers);

- Health care professionals;

- Payers (e.g., health insurance plans); and

- Government.

The basic structure, content, and operation of databases will be the same to address the interests of the different audiences. The desired information and how it is presented will, however, vary by audience.

Supporting the needs of the *public* is particularly important. It cannot be assumed that the public supports this work; in fact, there may be a certain amount of public distrust and hostility. The critical need is to develop and convey information that creates public trust and confidence. People must be assured that their privacy will be maintained. For the public, a focus on validity and utility is important; these considerations need to be conveyed in terms of the value, potential harms, and costs of a test for an individual and his or her family.

Professionals want to know about harms, benefits, and long-term outcomes, but they are also interested in payment issues. Both guidelines and patient-specific decision tools are needed to facilitate widespread use within the profession.

Patient electronic medical records must incorporate three components to be useful: the clinical information itself, the decision support tools, and a way to track how information is tailored for the patient. Language in the form of disease coding (e.g., SNOMED) is an important obstacle here. Lack of a central database—such as a master patient index—could be a limitation of electronic medical records, because this lack would preclude searching across institutions or across states.

One major task of the *payer* is adjudication of insurance claims. Payers make decisions about individual patients, but they do so on the basis of population-based studies. Therefore, payers will want to know about several things that relate to any gene-based intervention: clinical utility; long- and short-term patient outcomes; and whether a test is accepted by patients, by advocacy and professional groups, and by other payers.

The *government and policymakers* must take a broad view and consider the interests of all individuals and groups in a society. For setting policy on gene-based applications, data are needed on utilization trends, utilization patterns and potential disparities, appropriateness of use, impact on health care practice and costs, benefits and risks, and long-term health outcomes and cost-effectiveness.

Proposed Characteristics of an Optimal Database or System of Databases

The desired characteristics of a database system for monitoring utilization and outcomes of gene-based applications will depend on the specific purposes for which the system would be used. A comprehensive system that could provide on-going data on all aspects of gene-based testing—from monitoring test utilization to assessing impact on health care interventions and patient outcomes—is not likely to be feasible in the US health care system in the near future.

Linking databases, however, to acquire some of the desired information for a given test may be possible. Any monitoring system should build as much as possible on existing health systems databases. For example, newborn screening databases have already been linked to birth certificates in some states. A great deal of information could be obtained if these databases were further linked to information from the public health genetics programs that coordinate followup, and to state birth defects and developmental disabilities surveillance programs.

The key features and considerations for any database or system of databases for monitoring utilization and outcomes of gene-based applications include the following:

- The database system must contain high-quality genetic testing, clinical, and patient data that can be linked at the individual level.

- Data quality and standards are paramount considerations. Key concerns include:
  – Standard coding of genetic tests (more specific, include test result),
  – Common standards for interoperability (linkage, harmonization), and
  – Inclusion of family history data.

- The database system must address issues of incongruity between claims processing and research needs.

- The different needs of clinical (individual-level) monitoring vs. population-based surveillance should be kept in mind.

- Considering the audience for summary information is crucial. As discussed above, they include
  - The public, including patients and their families,
  - Health care providers,
  - Insurance carriers and other payers, and
  - Government agencies.

Although each audience may need different summary data to guide decisions, all their needs require a monitoring system based on individual-level data with laboratory test results linked to health outcome data.

- Developers and users of such databases must consider the implications of an individual's genetic test results for his or her family.

- Incentives for development or enhancement of databases will be important. They include:
  - pay-for-performance incentives (e.g., "coverage with evidence" or "phased reimbursement"),
  - Involvement of trusted third parties (e.g., for database management and administration), and
  - Incentives to industry, such as uses for postmarketing surveillance.

An all-encompassing system as envisioned above may be problematic at any time; certainly it is not currently available. Covering all the types of genetic testing that occur today would be an extremely complex undertaking. Any database system will need to be flexible and extensible to accommodate the addition of new tests and types of tests and be able to harmonize across databases through the use of common data elements and common vocabulary.

As noted earlier, workshop participants and the project's key informants provided important insights into steps needed to establish a surveillance system. Those who supported the general idea recommended using the existing pieces if possible; they often emphasized the cost of setting up a system yet cautioned that existing data acquisition might have its own challenges. An alternative view was that the current approach used now—individual studies—may be more appropriate for most of the questions researchers and decision makers are interested in.

An intermediate position may be to focus on data systems that could address the three broad purposes that guided our assessment: (1) test utilization and knowledge of test results; (2) impact on treatment choices and short-term clinical outcomes; and (3) health and long-term outcomes. The prospects, potential databases, and considerations for establishing databases for these three purposes were a focus of the discussions by the three workgroups at the project workshop. Several issues emphasized by one or another of these work groups are actually cross-cutting and would apply to databases in general.

Tables 8-10 outline key considerations on the development of relevant databases that emerged from the discussions of the three work groups at the invitational workshop. Respectively, they deal with use of tests and knowledge of test results, impact on treatment choices and short-term outcomes, and then long-term outcomes. Each table first summarizes current databases and database needs and then highlights possible future directions.

Table 8. Test utilization and knowledge of test results

- Summary of current databases and database needs

    o Monitoring utilization requires medical/health records linked to laboratory data at patient level

    o Data are not generally available

        - May be available for large managed care organizations, the Department of Veterans Affairs, or the Department of Defense
        - May be available at state level for newborn screening and other public health applications

    o Requirements for establishing a database system or modifying existing database systems

        - Standardization of terminology across databases
        - Improved health records with better data on family history
        - Detailed information about disease and pathology
        - More specific codes regarding genetic testing
        - Common standards for interoperability
        - Sustainability over time (e.g., anticipated profusion in new tests and testing technologies, changes in databases and data standards)

- Possible future directions

    o Design one or more pilot projects focusing on specific tests and diseases

        - The necessary steps could include forming a multidisciplinary group of experts, considering what tests are available and defining needs, and defining disease-specific parameters that may influence what information is required in the clinical history.

Table 9. Impact on treatment choices and short-term clinical outcomes

- Summary of current databases and database needs

    o Relationship between genetic tests, associated databases, and clinical outcomes is fractured.

    o Major need is for vocabulary standardization to allow linkage of existing systems and data sharing

    o Difficulty of harmonization is compounded by the existence of multiple systems and data warehouses

    o These issues limit the ability to generate effectiveness and outcomes related data.

    o The situation with databases on gene-based applications is reflective of the current overall state of health information technology (IT)

- Possible future directions

    o Facilitate information exchange through standardization (identify and focus on successful models)

    o Consider incentives for industry

        - Sharing data may ultimately enhance a firm's product uptake
        - Vendor education on data collection

    o Address the incongruity between information needed to process claims or have a functional electronic health record and research data

    o Begin by conducting pilot programs

        - Genetic testing databases provide an opportunity for a health IT success story
        - Explore partnerships (e.g., RHIOs)[72]
        - Genetic testing for warfarin therapy may be a useful model (there are few suppliers and few diseases for which it is needed, but 2 million people use the drug every year)

IT, information technology; RHIOs, Regional Health Information Organizations.

Table 10. Health and long-term outcomes

- Summary of current databases and database needs
    - o High premium on high quality patient data, including patient self-report (health status and functioning and quality of life)
    - o Build on existing health systems databases
        - Any administrative database that can link laboratory results with claims data should be useful for determining outcomes
    - o Data quality and standards
        - Standard coding of tests including results
        - Need unique identifier to link across databases
        - Need coded family history data
    - o Followup of individuals for health outcomes is a necessity, but may be problematic
    - o Clinical vs. population-based surveillance will require different data sources
        - Clinical surveillance will cover those with insurance and access to medical care (e.g., administrative, claims, health records data)
        - Population surveillance would require data from broad-based surveys including individuals without health insurance or accessible health data (e.g., self-pay to avoid disclosing health information, direct-to-consumer testing)
        - Cut across socioeconomic status categories
        - Detect disparities
    - o Consider effects on family
    - o Questions to consider
        - Will these issues get folded into the "pay-for-performance" movement (health care payers are increasingly requiring evidence of clinical utility before they will pay for genetic tests [5])?
        - What does the future hold with proliferation of genetic tests and new technologies (e.g., genome-wide scans)?
- Possible future directions
    - o Develop international coding standards for genetic testing data
    - o Standards need to be developed that will allow interoperability of health-related databases
    - o Support the development of statistical and health services research methodology to deal with the complexity of genetic data
    - o Conduct pilot studies to begin to develop databases for evaluating health outcomes
        - BRCA1/2 screening of asymptomatic women for breast/ovarian cancer risk may be a good model (Myriad Genetics is the sole provider of the test and long-term outcomes data should be available from various clinical trials)
        - Quality of life data are hard to evaluate for some of the aggressive interventions used for breast cancer treatment and prevention; adverse events resulting from experimental drugs such as tamoxifen may be difficult to abstract
        - The Breast Cancer Family Registry and the Cancer Genetic Network may provide useful data, supplemented by the Cancer Research Network and Medicare may be able to provide data if screening has been covered.

The next sections synthesize the findings of our assessment and the workshop deliberations in considering ways to achieve the three broad goals of such databases outlined in Tables 8-10.

Test Utilization and Knowledge of Test Results

Monitor Utilization of Genetic Tests

Monitoring use of tests is probably the most straightforward goal and one that potentially could be addressed using existing databases. To determine what information is available in current databases, the best source is probably the laboratory performing the test, although they will often not know why a test was done, how many people in the population require the test (i.e., the denominator in the population prevalence equation), or how many other laboratories offer the same test. Nevertheless, even when a test is performed in relatively few laboratories determining with certainty what information is currently available that could be used to address specific research questions may still not be possible. The main reason is that critical elements in the databases may be poorly defined. If laboratory results were linked to patient medical records—perhaps using electronic medical records—determining what information is available would be more feasible.

Potential data sources include genetic laboratory databases, large commercial laboratory databases, databases of large health plans (e.g., MCOs, large employer health plans), Medicare and Medicaid databases and the databases of the Department of Veterans Affairs (specifically, the Veterans Health Administration), and the Department of Defense (DoD). To monitor utilization trends, these databases should collect data on an ongoing ("real time") basis, cover large, defined populations, and have standardized coding schemes for specific genetic tests (e.g., type of test, reason for test, and test result).

None of the databases evaluated, however, could be used without additional research and development to address limitations in coding (e.g., test type, indication, and result) and limitations of access and integration. No one database exists that could provide national-level data on utilization trends. Thus, the absolute magnitude of testing volume at the national level could only be estimated. Relative trends in testing could be accessed through monitoring in several data sources that provide reliable data for specific regions or defined populations.

Knowledge of Test Results and Appropriateness of and Access to Use of Genetic Tests and Gene-Based Interventions

Assessing appropriateness of gene-based tests and related interventions, patient and provider knowledge about the interventions, information shared and actions taken in response to tests, and possible gaps and disparities in the utilization of gene-based interventions is likely to be difficult using available health care databases. Additional data gathering from individuals and health care providers through surveys will probably be required. Through longitudinally linked databases of large providers (e.g., KP HealthConnect), some initial assessments may be possible to determine whether patients identified with certain high-risk genetic traits are receiving the recommended therapies.

Similarly, differences in utilization and possible disparities may also be tracked through large health care databases with data on patient characteristics (e.g., age, sex, race). For example, the Central New York Regional Perinatal Data System (RPDS) is a population-based birth registry that includes information on demographics, antepartum, intrapartum, and postpartum care, and

neonatal outcomes; similar data systems are available for other regions of New York. The RPDS and any similar databases could be used to monitor utilization and outcomes of pregnancy associated gene-based tests. None of these databases, however, provides national-level data.

Information on knowledge and actions taken in response to testing could be obtained through special surveys or perhaps by adding questions to ongoing large surveys, such as the Pregnancy Risk Assessment Monitoring System (PRAMS), Medicare Health Outcomes Survey, and the National Ambulatory Medical Care Survey (NAMCS).

Impact on Treatment Choices and Short-Term Clinical Outcomes

This goal may be achievable in a limited fashion using existing databases. The critical requirement is coding and database system standards to be able to link specific gene-based tests to treatments and immediate results of treatment. Databases that potentially could be used for this purpose include those that can link information on both diagnostic testing and clinical treatment or management. These might include administrative databases (e.g., beneficiary claims from the Centers for Medicare & Medicaid Services [CMS], large insurance data systems, MCOs, and other health plans), the NAMCS, selected ongoing research cohorts (e.g., Marshfield Clinic, Kaiser Program on Genes, Environment and Health, Cancer Biomedical Informatics Grid), and certain registries (e.g., SEER, National Breast and Colon Cancer Family Registries).

All current potential data sources, however, will require some modification or additional data collection. This might include medical record review to obtain additional clinical details and to validate database information, addition of supplemental questions to ongoing surveys, and similar steps.

The development and operation of databases for gene-based treatment are similar to the challenges in databases for other purposes, and are reflective of the current state of health IT in general. Creating and maintaining databases for gene-based interventions are relatively more focused undertakings than broader IT initiatives. Successful development of this specific area could feasibly provide a success story in the advancement of databases and health IT for medical care research and evaluation.

Some steps that could be taken in this regard would be to explore incentives to encourage commercial laboratories to make their data accessible (e.g., the possibility that demonstrating the positive health impact and cost-effectiveness of a particular test could lead to increased utilization of the test). Developing partnerships among various "owners" of databases, along the line of the model offered by Regional Health Information Organizations (RHIO),[72] may also be fruitful. Related efforts in health IT being considered by SACGHS and AHIC should also facilitate the development and appropriate use of interoperable patient-level data for research and clinical decision making.

Health and Long-Term Outcomes

Monitoring outcomes of gene-based interventions over the long run is the most challenging undertaking. Outcomes for specific gene-based interventions would have to be defined. For

outcomes associated with a distinct disease that has a clearly established diagnostic code, such as breast cancer, conducting longitudinal linkages in databases such as those of large health systems might be possible. Outcomes that are based on quality of life or functional status would require establishing longitudinal surveys in which individuals are interviewed using appropriate standardized questionnaires. Survey data would also be required for "population-based" studies that seek to capture information from all individuals, not just those with documented health care encounters.

If a highly effective treatment should be developed and widely implemented for a common gene-based disease, determining the impact of the intervention in ecological data (e.g., national or state hospital discharge data, mortality statistics) might be possible. For example, these types of ecological analyses have been very powerful in demonstrating the dramatic impact of immunization programs.

To evaluate outcomes of gene-based applications, representativeness may not be as crucial a requirement as it is for monitoring test utilization. Thus, more restricted databases that do not cover a large or diverse population may still be useful for evaluation of outcomes, provided that the data have internal validity. Some of the databases that could potentially be used to monitor and evaluate outcomes of gene-based testing and related treatments include the following: the Newborn Screening Information System, the NAMCS, MCO data, the Marshfield and Kaiser research cohorts, the Cancer Biomedical Informatics Grid, SEER, and the National Breast and Colon Cancer Family Registries. The Breast Cancer Family Registry and the Cancer Genetic Network may be particularly valuable sources for outcomes of cancer.

Other options for research (more so than monitoring) include the entire Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) Network itself at the Agency for Healthcare Research and Quality (AHRQ) (http://effectivehealthcare.ahrq.gov/index.cfm). The Centers for Education and Research on Therapeutics (CERTs) program at AHRQ, mandated by the US Congress almost a decade ago and affiliated with AHRQ's Effective Health Care program, is another resource. CERTs centers are not focused on genomics, but at least three centers have an interest in genetic issues and may be useful for studies of outcomes of gene-based applications and other genetic testing questions. AHRQ's programs in Accelerating Change and Transformation in Organizations and Networks (ACTION) and the two pilot DEcIDE projects on Distributed Research Network in Therapeutics (for which specifications are being developed by HMO Research Network and the University of Colorado) may provide similar opportunities.

Recommendations for Further Research and Development

A major focus of our project was to develop recommendations for additional steps, including research, to enhance the utility of current databases or advance the development of new databases or systems of databases to monitor utilization and impacts of gene-based tests and interventions in the health care system. We devoted considerable attention to this need at the workshop. Throughout this chapter, we have discussed the problems and limitations of current databases and data collection efforts that impair our ability to monitor and evaluate gene-based testing, interventions, and outcomes. We also suggested potential solutions to many of the specific problems identified. In summary, we make the following recommendations for steps that can be taken to improve the databases and other information systems to enable improved

monitoring and research on utilization and outcomes of gene-based tests and related interventions:

- Improve the coding of gene-based tests in many of the relevant databases so that the test type, reason for the test, and test results could be readily determined.
  – Relevant agencies and offices of the Department of Health and Human Services, perhaps in conjunction with the World Health Organization, should consider convening a working group to establish international standards for coding genetic tests.

- Develop or adopt standards for the proper collection and storage of data from genetic testing laboratories for the archiving of the tests performed and facilitating interoperability between databases.
  – Consideration should be given to issues that will allow linkage with clinical databases and electronic health records.
  – Include minimal data requirements for databases on gene-based applications for screening and diagnosis (e.g., age, race and ethnicity, family history, relevant clinical details).
  – Consider how to link between existing databases using data mining tools and search engines.
  – Address privacy and confidentiality concerns that may limit data sharing, taking into account the recommendations of relevant advisory groups such as SACGHS, AHIC, and the National Committee for Vital and Health Statistics.

- Explore the possibility of adding questions to ongoing surveys or developing new surveys to monitor the availability of genetic testing centers, adequate counseling, and barriers to accessing counseling services, as well as direct-to-consumer testing.

- Consider establishing a survey of genetic testing laboratories similar to the NAMCS for medical clinics and the National Hospital Discharge Survey for hospitals.

- Develop pilot studies for a small set of diseases and tests (e.g., BRCA1/2 screening, newborn screening, genetic testing for warfarin therapy).
  – Establish a work group to develop the pilot studies (pick diseases, tests, disease-specific parameters).
  – Build on the experiences of other projects, such as those supported by AHRQ's Effective Health Care program, including the DEcIDE Research Network, the CERTs program, the ACTION program, and the Distributed Network for Ambulatory Research in Therapeutics.
  – Utilize existing disease registries to the extent possible to study outcomes of testing and related interventions.

- Garner support and trust for databases and research on gene-based tests and related applications.

- Public trust will require public and provider education not only about the various regulatory and technological protections built into the databases, but more importantly about the value of the databases in improving personal health care.
- Education and communication efforts should be geared toward health care practitioners and students, public health workers, and consumers.
- Volunteer patient support groups could potentially play an important role in facilitating public support.

Conclusion

Recent advances in genetic science, especially the sequencing of the human genome, have generated great expectations of a revolution in "personalized medicine" in which interventions to prevent and treat disease will be tailored to each individual's genetic composition. Many gene-based tests have been developed and many more are under development, including genome-wide scans that can already measure hundreds of thousands of genetic markers. The development of gene-based tests, however, is outpacing the evaluation of their validity and utility. Many tests have entered clinical practice, and some are being marketed directly to consumers, without adequate knowledge of how well they perform in identifying or predicting risks of particular health conditions or the outcomes of treatment decisions based on the genetic test results. Thus, a need exists at the public health level to be able to monitor the utilization of gene-based tests to be able to determine trends in use, appropriateness of use, and potential disparities in utilization. Health care policymakers, providers, and payers need data on how specific genetic tests and related interventions impact short- and long-term health outcomes, including information on cost-effectiveness of specific tests and any subsequent interventions. For purposes of public health surveillance, databases are not currently available that would enable national-level monitoring of gene-based testing. For clinical utility assessment, current health databases and other information sources are adequate to allow some health services research to address certain specific questions. Health services research in this area could be much advanced, however, if existing laboratory, clinical, health plan, and disease registry databases could be linked to allow large and efficient studies of outcomes of genetic tests and gene-based interventions.

A number of challenges will have to be addressed before the ideal of being able to compile and link data from existing health databases and surveys for public health surveillance and health services research can be realized. These include developing standard codes for genetic tests and database architecture standards to allow interoperability between databases; and addressing concerns about privacy and confidentiality, as well as proprietary and regulatory barriers that inhibit sharing of data. The needed technical advances, such as standards to facilitate database interoperability, are beginning to be addressed by national initiatives to improve health information technology and promote personalized health care. Ultimately, the development and operation of databases and other data collection efforts to monitor utilization and outcomes of gene-based tests and related interventions will require public trust and support. Public support will require confidence in the security built into the databases or information systems to protect personal health information and an understanding of the value of the databases in improving personal health care.

References

1. Secretary's Advisory Committee on Genetic Testing. Highlights of the Fifth Meeting of the Secretary's Advisory Committee on Genetic Testing, June 5–7, 2000. Secretary's Advisory Committee on Genetic Testing; 2000; Washington, DC. Available from: http://www4.od.nih.gov/oba/sacgt/highlights/highlights62000.pdf; 2000.

2. GeneTests. [cited January 16, 2007]; Available from: http://www.genetests.org/

3. Alsever J. The gene screen. October 3, 2006 [cited; Available from: Accessed January 16, 2007, from http://money.cnn.com/2006/09/26/magazines/business2/8387104.biz2/index.htm?postversion=2006092706

4. U.S. Preventive Services Task Force (USPSTF). http://www.ahrq.gov/clinic/uspstf/uspsbrgen.htm.

5. U.S. System of Oversight of Genetic Testing: A Response to the Charge of the Secretary of Health and Human Services. Report of the Secretary's Advisory Committee on Genetics, Health, and Society. April 2008. Available from: http://www4.od.nih.gov/oba/SACGHS/reports/SACGHS_oversight_report.pdf.

6. Mouchawar J, Hensley-Alford S, Laurion S, Ellis J, Kulchak-Rahm A, Finucane ML, et al. Impact of direct-to-consumer advertising for hereditary breast cancer testing on genetic services at a managed care organization: a naturally-occurring experiment. Genet Med. 2005 Mar;7(3):191-7.

7. Centers for Disease Control and Prevention (CDC). Genetic testing for breast and ovarian cancer susceptibility: evaluating direct-to-consumer marketing--Atlanta, Denver, Raleigh-Durham, and Seattle, 2003. MMWR Morb Mortal Wkly Rep. 2004;53(27):603-6.

8. Hunter DJ, Khoury MJ, Drazen JM. Letting the genome out of the bottle--will we get our wish? N Engl J Med. 2008;358:105-7.

9. Whitehead NS, Rasmussen SA, Cox S, Posner SF. Prevalence and predictors of receipt of prenatal information about genetic screening. Prenat Diagn. 2006;26(10):944-50.

10. University of Kansas Medical Center. Genetic Testing Labs in United States. [cited July 10, 2007]; Available from: http://www.kumc.edu/gec/prof/labs.html

11. National Newborn Screening and Genetics Resource Center. National Newborn Screening Information System. [cited June 29, 2007]; Available from: http://www2.uthscsa.edu/nnsis/

12. U.S. National Library of Medicine. Genetics Home Reference. 2007/08/24/ [cited; Available from: http://ghr.nlm.nih.gov/ghr/search/show/conditions?query=%22newborn+screening%22

13. Kaiser Permanente. Health Care Marketplace | Kaiser Permanente Launches Long-Term Project To Collect Genetic Data on Members. [cited February 15, 2007]; Available from: http://www.kaisernetwork.org/daily_reports/rep_index.cfm?hint=3&DR_ID=42962

14. National Institute of Mental Health (NIMH). Human Genetic Initiative. [cited; Available from: http://zork.wustl.edu/nimh/NIMH_initiative/NIMH_initiative_link.html

15. Autism Genetic Resource Exchange. Program Description. [cited July 27, 2007]; Available from: http://www.agre.org/program/descr.cfm?do=program

16. Mountcastle-Shah E, Holtzman NA. Primary care physicians' perceptions of barriers to genetic testing and their willingness to participate in research. Am J Med Genet. 2000;94(5):409-16.

17. Batra S, Valdimarsdottir H, McGovern M, Itzkowitz S, Brown K. Awareness of genetic testing for colorectal cancer predisposition among specialists in gastroenterology. Am J Gastroenterol. 2002;97(3):729-33.

18. Chase GA, Geller G, Havstad SL, Holtzman NA, Bassett SS. Physicians' propensity to offer genetic testing for Alzheimer's disease: results from a survey. Genet Med. 2002;4(4):297-303.

19. Cho MK, Sankar P, Wolpe PR, Godmilow L. Commercialization of BRCA1/2 testing: practitioner awareness and use of a new genetic test. Am J Med Genet. 1999;83(3):157-63.

20. Wideroff L, Freedman AN, Olson L, Klabunde CN, Davis W, Srinath KP, et al. Physician use of genetic testing for cancer susceptibility: results of a national survey. Cancer Epidemiol Biomarkers Prev. 2003;12(4):295-303.

21. Koil CE, Everett JN, Hoechstetter L, Ricer RE, Huelsman KM. Differences in physician referral practices and attitudes regarding hereditary breast cancer by clinical practice location. Genet Med. 2003;5(5):364-9.

22. Sifri R, Myers R, Hyslop T, Turner B, Cocroft J, Rothermel T, et al. Use of cancer susceptibility testing among primary care physicians. Clin Genet. 2003;64(4):355-60.

23. Friedman LC, Cooper HP, Webb JA, Weinberg AD, Plon SE. Primary care physicians' attitudes and practices regarding cancer genetics: a comparison of 2001 with 1996 survey results. J Cancer Educ. 2003;18(2):91-4.

24. Morgan MA, Driscoll DA, Mennuti MT, Schulkin J. Practice patterns of obstetrician-gynecologists regarding preconception and prenatal screening for cystic fibrosis. Genet Med. 2004;6(5):450-5.

25. Cleary-Goldman J, Morgan MA, Malone FD, Robinson JN, D'Alton ME, Schulkin J. Screening for Down syndrome: practice patterns and knowledge of obstetricians and gynecologists. Obstet Gynecol. 2006;107(1):11-7.

26. Wilkins-Haug L, Erickson K, Hill L, Power M, Holzman GB, Schulkin J. Obstetrician-gynecologists' opinions and attitudes on the role of genetics in women's health. J Womens Health Gend Based Med. 2000;9(8):873-9.

27. Freedman AN, Wideroff L, Olson L, Davis W, Klabunde C, Srinath KP, et al. US physicians' attitudes toward genetic testing for cancer susceptibility. Am J Med Genet A. 2003;120(1):63-71.

28. National Center for Health S. National Ambulatory Medical Care Survey. 2007 [cited 2007/08/22/]; Available from: http://www.cdc.gov/nchs/about/major/ahcd/namcsdes.htm

29. Centers for Disease Control and Prevention (CDC). National Hospital Ambulatory Medical Care Survey (NHAMCS). [cited 9/13/2007]; Available from: http://www.cdc.gov/nchs/about/major/ahcd/nhamcsds.htm

30. National Center for Health S. National Health Interview Survey. 2007 [cited 2007/08/22/]; Available from: http://www.cdc.gov/nchs/about/major/nhis/hisdesc.htm

31. National Center for Health S. National Health and Nutrition Examination Survey, 2005-2006 Overview. 2007 [cited August 22, 2007]; Available from: http://www.cdc.gov/nchs/about/major/nhanes/intro_mec.htm; http://www.cdc.gov/nchs/data/nhanes/OverviewBrochureEnglish_May05.pdf

32. National Center for Chronic Disease P, Health P. Behavioral Risk Factor Surveillance System; About the BRFSS. 2007 [cited 2007/08/22/]; Available from: http://www.cdc.gov/brfss/about.htm

33. Summerton N, Garrood PV. The family history in family practice: a questionnaire study. Fam Pract. 1997;14(4):285-8.

34. Cheschier N. ACOG practice bulletin. Neural tube defects. Number 44, July 2003. (Replaces committee opinion number 252, March 2001). Int J Gynaecol Obstet. 2003;83(1):123-33.

35. Wojtowycz MA, Dye TD, ubry RH. Disparities in the delivery of genetic screening tests associated with race. Abstr Book Assoc Health Serv Res Meet. 1997;14:98-9.

36. Strom CM, Crossley B, Redman JB, Buller A, Quan F, Peng M, et al. Molecular testing for Fragile X Syndrome: lessons learned from 119,232 tests performed in a clinical laboratory. Genet Med. 2007;9(1):46-51.

37. Kievit W, de Bruin JH, Adang EM, Ligtenberg MJ, Nagengast FM, van Krieken JH, et al. Current clinical selection strategies for identification of hereditary non-polyposis colorectal cancer families are inadequate: a meta-analysis. Clin Genet. 2004;65(4):308-16.

38. Lynch HT, Riley BD, Weissman SM, Coronel SM, Kinarsky Y, Lynch JF, et al. Hereditary nonpolyposis colorectal carcinoma (HNPCC) and HNPCC-like families: Problems in diagnosis, surveillance, and management. Cancer. 2004;100(1):53-64.

39. Bonis PA, Trikalinos TA, Chung M, Chew P, Ip S, DeVine D, 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.

40. American College of Medical Genetics' Newborn Screening Expert Group. Newborn screening: Toward a uniform screening panel and system. Genet Med. 2006;8:1S-52S; http://www.acmg.net/resources/policies/NBS/NBS-sections.htm.

41. Strom CM, Crossley B, Redman JB, Buller A, Quan F, Peng M, et al. Cystic fibrosis screening: lessons learned from the first 320,000 patients. Genet Med. 2004;6(3):136-40.

42. Strom CM, Crossley B, Redman JB, Quan F, Buller A, McGinniss MJ, et al. Molecular screening for diseases frequent in Ashkenazi Jews: lessons learned from more than 100,000 tests performed in a commercial laboratory. Genet Med. 2004;6(3):145-52.

43. Muller PR, Cocciolone R, Haan EA, Wilkinson C, Scott H, Sage L, et al. Trends in state/population-based Down syndrome screening and invasive prenatal testing with the introduction of first-trimester combined Down syndrome screening, South Australia, 1995-2005. Am J Obstet Gynecol. 2007;196(4):315-7.

44. O'Leary P, Breheny N, Dickinson JE, Bower C, Goldblatt J, Hewitt B, et al. First-trimester combined screening for Down syndrome and other fetal anomalies. Obstet Gynecol. 2006;107(4):869-76.

45. Benn PA, Ying J. Preliminary estimate for the second-trimester maternal serum screening detection rate of the 45,X karyotype using alpha-fetoprotein, unconjugated estriol and human chorionic gonadotropin. J Matern Fetal Neonatal Med. 2004;15(3):160-6.

46. National Genetics Institute. National Genetics Institute. [cited; Available from: http://www.ngi.com/

47. Quest Diagnostics. Diagnostic testing and services. [cited July 27, 2007]; Available from: http://www.questdiagnostics.com/brand/business/b_bus_lab_index.html

48. National Cancer Institute. caBIG TM : The power of connection. 2007 [cited; Available from: http://cabig.cancer.gov/connection/vision.asp

49. Therrell BL, Jr. Data integration and warehousing: coordination between newborn screening and related public health programs. Southeast Asian J Trop Med Public Health. 2003;34 Suppl 3:63-8.

50. National Cancer Institute. SEER Data, 1973-2004. [cited July 31, 2007]; Available from: http://seer.cancer.gov/data/

51. Gringras P. Choice of medical investigations for developmental delay: a questionnaire survey. Child Care Health Dev. 1998;24(4):267-76.

52. Rauch A, Hoyer J, Guth S, Zweier C, Kraus C, Becker C, et al. Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. Am J Med Genet A. 2006;140(19):2063-74.

53. Sequist LV, Joshi VA, Janne PA, Muzikansky A, Fidias P, Meyerson M, et al. Response to treatment and survival of patients with non-small cell lung cancer undergoing somatic EGFR mutation testing. Oncologist. 2007;12(1):90-8.

54. National Cancer Institute. NCI Research Resources. [cited July 31, 2007]; Available from: http://resresources.nci.nih.gov/database.cfm?id=1241

55. Berk T, Cohen Z, Bapat B, Gallinger S. Negative genetic test result in familial adenomatous polyposis: clinical screening implications. Dis Colon Rectum. 1999;42(3):307-10.

56. Tinley ST, Houfek J, Watson P, Wenzel L, Clark MB, Coughlin S, et al. Screening adherence in BRCA1/2 families is associated with primary physicians' behavior. Am J Med Genet A. 2004;125(1):5-11.

57. Kolibianakis E, Osmanagaoglu K, De CL, Camus M, Bonduelle M, Liebaers I, et al. Prenatal genetic testing by amniocentesis appears to result in a lower risk of fetal loss than chorionic villus sampling in singleton pregnancies achieved by intracytoplasmic sperm injection. Fertil Steril. 2003;79(2):374-8.

58. Moller P, Reis MM, Evans G, Vasen H, Haites N, Anderson E, et al. Efficacy of early diagnosis and treatment in women with a family history of breast cancer. European Familial Breast Cancer Collaborative Group. Dis Markers. 1999;15(1-3):179-86.

59. van Roosmalen MS, Stalmeier PF, Verhoef LC, Hoekstra-Weebers JE, Oosterwijk JC, Hoogerbrugge N, et al. Impact of BRCA1/2 testing and disclosure of a positive test result on women affected and unaffected with breast or ovarian cancer. Am J Med Genet A. 2004;124 (4):346-55.

60. Gritz ER, Peterson SK, Vernon SW, Marani SK, Baile WF, Watts BG, et al. Psychological impact of genetic testing for hereditary nonpolyposis colorectal cancer. J Clin Oncol. 2005 Mar 20;23(9):1902-10.

61. Murakami Y, Okamura H, Sugano K, Yoshida T, Kazuma K, Akechi T, et al. Psychologic distress after disclosure of genetic test results regarding hereditary nonpolyposis colorectal carcinoma. Cancer. 2004 Jul 15;101(2):395-403.

62. Keller M, Jost R, Haunstetter CM, Kienle P, Knaebel HP, Gebert J, et al. Comprehensive genetic counseling for families at risk for HNPCC: impact on distress and perceptions. Genet Test. 2002 Winter;6(4):291-302.

63. Benatti P, Roncucci L, Ganazzi D, Percesepe A, Di Gregorio C, Pedroni M, et al. Clinical and biologic heterogeneity of hereditary nonpolyposis colorectal cancer. Int J Cancer. 2001 Sep 20;95(5):323-8.

64. Fujita S, Moriya Y, Sugihara K, Akasu T, Ushio K. Prognosis of hereditary nonpolyposis colorectal cancer (HNPCC) and the role of Japanese criteria for HNPCC. Jpn J Clin Oncol. 1996 Oct;26(5):351-5.

65. Marchionni L, Wilson RF, Marinopoulos SS, Wolff AC, Parmigiani G, Bass EB, et al. Impact of Gene Expression Profiling Tests on Breast Cancer Outcomes. Rockville, Maryland: Agency for Healthcare Research and Quality; 2008 January 2008.

66. Umans-Eckenhausen MA, Defesche JC, van Dam MJ, Kastelein JJ. Long-term compliance with lipid-lowering medication after genetic screening for familial hypercholesterolemia. Arch Intern Med. 2003;163(1):65-8.

67. American Health Information Community. [cited 10/22/07;]. Available from: http://www.hhs. gov/healthit/community/background/

68. Microsoft HealthVault beta. [cited 10/22/07]; Available from: http://www.healthvault.com/

69. National Health Information Network. [cited 11/08/07]; Available from: http://www.ansi.org/ standards_activities/standards_boards_panels/his b/hitsp.aspx?menuid=3

70. Public Health Information Network. [cited 11/08/07]; Available from: http://www.cdc.gov/ phin/about.html

71. National Committee on Vital and Health Statistics (NCVHS). Letter to the Secretary: Recommendations regarding privacy and confidentiality in the National Health Information Network. National Committee on Vital and Health Statistics, Washington, DC. http://www.ncvhs.hhs.gov/060622l5.htm 2006.

72. Office of the National Coordinator for Health Information Technology: Goals of Strategic Framework. [cited 10-22-07]; Available from: http://www.hhs.gov/healthit/goals.html

Glossary

ACOG	American College of Obstetricians and Gynecologists
ACTION	Accelerating Change and Transformation in Organizations and Networks
AHIC	American Health Information Community
AHRQ	Agency for Healthcare Research and Quality
APOE	Apolipoprotein E
BCP	Breast Cancer Profiling
BRCA1/2	Breast cancer 1/2
BRFSS	Behavioral Risk Factor Surveillance System
CA125	Cancer antigen 125
caBIG™	Cancer Biomedical Informatics Grid™
CanCORS	Cancer Care Outcomes Research and Surveillance Consortium
CAP	College of American Pathologists
CDC	Centers for Disease Control and Prevention
CLIA	Clinical Laboratory Improvement Amendments of 1988
CMS	Centers for Medicare & Medicaid Services
CPT	Current Procedural Terminology
CRN	Cancer Research Network
CVS	Chorionic villus sampling
DEcIDE	Developing Evidence to Inform Decisions about Effectiveness
DNA	Deoxyribonucleic acid
DoD	Department of Defense
EGAPP	Evaluation of Genomic Applications in Practice and Prevention
EGFR	Epidermal growth factor receptor
FDA	Food and Drug Administration
FEHBP	Federal Employees Health Benefit Plan
FERPA	Family Educational Rights and Privacy Act
hCG	Human chorionic gonadotropin
HIPAA	Health Insurance Portability and Accountability Act of 1996
HL7	Health Level 7
HMO	Health maintenance organization
HNPCC	Hereditary nonpolyposis colorectal cancer
HRSA	Health Resources and Services Administration
ICD	International Classification of Diseases
ICD-9-CM	International Classification of Diseases (ninth version, clinical modification)
ID	Identification
IDEA	Individuals with Disabilities Education Act
IRB	Institutional Review Board
IT	Information technology
KP	Kaiser Permanente
KPNW	Kaiser Permanente Northwest
LOINC	Logical Observation Identifier Names and Codes
MA	Medicare Advantage
MCHB	Maternal and Child Health Bureau
MCO	Managed care organization

MSAFP	Maternal serum-alpha-fetoprotein
NAMCS	National Ambulatory Medical Care Survey
NCHS	National Center for Health Statistics
NCI	National Cancer Institute
NGI	National Genetics Institute
NHAMCS	National Hospital Ambulatory Medical Care Survey
NHDS	National Hospital Discharge Survey
NHIN	National Health Information Network
NIH	National Institutes of Health
NNSGRC	National Newborn Screening and Genetics Resource Center
NNSIS	National Newborn Screening Information System
PGRN	Pharmacogenetics Research Network
PHC WG	Personalized Healthcare Workgroup
PHIN	Public Health Information Network
PRAMS	Pregnancy Risk Assessment Monitoring System
RHIO	Regional health information organization
RNA	Ribonucleic acid
RPDS	Regional Perinatal Data System
RTI	RTI International
SACGHS	Secretary's Advisory Committee on Genetics, Health and Society
SACGT	Secretary's Advisory Committee on Genetic Testing
SEER	Surveillance Epidemiology and End Results
SNOMED	Systematized Nomenclature of Medicine
SNP	Single nucleotide polymorphism
TEP	Technical expert panel
UHC	United HealthCare
US	United States
USPSTF	US Preventive Services Task Force
VA	Veterans Administration
VistA	Veterans Health Information Systems and Technology Architecture

Appendixes

Appendix A. Methods

Data collection for this project comprised several related activities: reviewing both published and gray literature, conducting key informant interviews, reviewing genomics testing databases of several types, and, most critical, convening an invitational workshop to discuss issues relating to databases that could provide information on both use and outcomes of gene-based applications. The remainder of this chapter documents our methods for these activities. Before describing our procedures, however, we provide brief definitions of some key terms and concepts important for this work on genetics databases, and we describe our Technical Expert Panel (TEP).

Technical Expert Panel

In developing and evaluating a model approach for monitoring the utilization and outcomes of gene-based applications in the US healthcare system, we engaged experts from a variety of disciplines and organizations. To obtain such assistance throughout the project, the RTI Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) Center established a TEP to consult on the project methodology, to help identify additional experts to participate in the workshop, and to provide comments on our final report. We put forth nominations; experts approved by the Agency for Healthcare Research and Quality (AHRQ) and the Centers for Disease Control and Prevention (CDC) were invited to participate on the panel. Eight representatives with expertise in diverse areas (genetic epidemiology, health services research, healthcare, genetic laboratory testing, genetic test development, policy makers, and information technology) agreed to serve on our TEP.

The TEP's role was to review materials and provide guidance on the definitions used to frame the project and the inclusion and exclusion criteria for the literature reviews, suggest additional databases to be investigated, and nominate potential key informants to interview regarding the issues of utilizing current and future genetic testing databases. Three members were able to participate in the invitational workshop; several others served as peer reviewers of the draft final report.

Literature Review

We conducted two types of literature reviews, one of the peer-reviewed literature and the other of gray literature. Our approaches to these two related activities are described here.

Peer-Reviewed Literature

This targeted scientific literature review of the scientific literature was intended to supplement both our gray literature search and our key informant interviews. These latter two data sources served as the primary methods of identifying relevant databases, but the scientific literature review provided some preliminary information. This activity was not meant to be a comprehensive review of the literature on the use or outcomes of gene-based tests. With more

than 1,100 genetic tests in clinical use, and more than 15,000 publications on genetic screening, a complete review was beyond the scope of this project. Rather, the literature review was designed to identify:

- Existing information about the utilization and outcomes of gene-based tests and interventions.

- Databases and resources available and their strengths and limitations for assessing utilization and outcomes of gene-based applications.

- Information regarding the optimal characteristics of a database or information network that could be used to assess the utilization and outcomes of gene-based tests and interventions.

We searched only MEDLINE®, using terms shown in Table A-1 and inclusion/exclusion criteria noted in Table A-2. We used three separate search strategies to identify relevant articles (see Table A-1).

Search 1 was designed to identify routinely collected and maintained databases that have information on gene-based tests and interventions. After we reviewed and abstracted the articles relevant to this search, we reviewed the reference lists and full PubMed citations for the most relevant articles. We did not find any additional relevant articles within the specified time period. In addition, we searched using additional terms, such as "Information Systems" and "Infomatics," but did not identify any other relevant articles.

Table A-1. Medical subject headings (MeSH) and their definitions

MeSH Terms	MeSH Definition or Statement of Scope	Used in Search
Molecular Diagnostic Techniques	MOLECULAR BIOLOGY techniques used in the diagnosis of disease. Included are such techniques as IN SITU HYBRIDIZATION of chromosomes for CYTOGENETIC ANALYSIS; OLIGONUCLEOTIDE ARRAY SEQUENCE ANALYSIS of gene expression patterns in disease states; identification of pathogenic organisms by analysis of species-specific DNA sequences; and detection of mutations with POLYMERASE CHAIN REACTION.	1, 2, 3
Genetic Screening	Searching a population or individuals for persons possessing certain genotypes or karyotypes that: (1) are already associated with disease or predispose to disease; (2) may lead to disease in their descendants; or (3) produce other variations not known to be associated with disease. Genetic screening may be directed toward identifying phenotypic expression of genetic traits. It includes prenatal genetic screening.	1,2, 3
Physician Practice Patterns	Patterns of practice related to diagnosis and treatment as especially influenced by cost of the service requested and provided.	2
Treatment Outcomes	Evaluation undertaken to assess the results or consequences of management and procedures used in combating disease in order to determine the efficacy, effectiveness, safety, practicability, etc., of these interventions in individual cases or series.	3
Database	Organized collections of computer records, standardized in format and content, that are stored in any of a variety of computer-readable modes. They are the basic sets of data from which computer-readable files are created.	1
Evaluation studies	Studies determining the effectiveness or value of processes, personnel, and equipment, or the material on conducting such studies.	3

Search 2 was designed to identify information regarding physicians' use of gene-based tests and interventions. We used the following search strategy: (*Molecular Diagnostic Techniques OR genetic screening) AND Physician's Practice Patterns.*

Search 3 was designed to identify information regarding health outcomes of gene-based tests and interventions. Our initial search strategy was (*Molecular Diagnostic Techniques OR genetic screening) AND Evaluation Studies*. This strategy yielded 33 pages of citations, most of which were clearly irrelevant. We searched the database of MeSH terms to identify better terms for a new search strategy; the final search strategy was (*Molecular Diagnostic Techniques OR genetic screening) AND Treatment Outcomes.*

Search Results

Databases. Our search for databases with information on gene-based tests and interventions identified 327 citations. We excluded 314 citations during the abstract review and requested 13 full articles for review. We were unable to obtain one article. Of the 12 articles reviewed, we excluded seven and abstracted data from five.

Use. Our search for information on utilization of gene-based tests and interventions identified 41 citations. We excluded 18 articles that were commentaries or reviews or were otherwise irrelevant and requested articles for the remaining 23 citations. After the review of the full articles, we excluded six as not relevant to this report and determined that four articles were more relevant for the treatment outcomes search. We abstracted 16 articles that provided information on the utilization of genetic tests and physician attitudes regarding genetic testing, including three articles identified during the database search.

Outcomes. Our search for information on outcomes of gene-based tests and interventions identified 122 citations. We excluded 103 articles that were commentaries or reviews or otherwise irrelevant, and requested articles for the remaining 19 citations, of which we were able to obtain 17. After review of these 17 articles, we determined that six were not relevant to this report. We abstracted data from 11 articles that provided information on outcomes of genetic tests.

Gray Literature

The sources for gray literature consisted of unpublished reports, news briefings, press conferences, web pages, and other information available electronically on the Internet. We used several Internet search engines, including Google and its relevant options (e.g., Google News, Google Scholar), Yahoo, and Ask.com. We entered the following search terms: "genetic testing data," "genetic testing database," "genetic testing surveillance," "genetic screening files/data," "genetic test database," "gene test," "data files health US," and "gene-based test data."

We followed various links on the web pages, identified through keyword search, when more specific information appeared to be available. In addition, if certain data systems were referred to in the key informant interviews (see below), then we also searched for these using Google search.

We then appraised the documents identified through these searches for their relevance to the project. A majority of the documents addressed various aspects of genetic testing without referring to data or files containing data on genetic testing. In other cases, health information databases were referred to without a reference to genetic testing. As a final step, we included only those documents containing information about data regarding genetic testing in this review.

Table A-2. Literature review inclusion and exclusion criteria

Category	Criteria
Populations	Human, all ages
Geography and Setting	All countries; all settings and sites
Publication date	1985 to May 15, 2007
Publication types	All except
	Searches 1 and 3: Research studies of genetic variants and disease*
	Search 2: Articles (e.g., policy discussions) that did not report data about use of gene-based tests*
Language	English only

*Excluded at the title and abstract review stage.

Key Informant Interviews

RTI, with inputs from AHRQ and CDC, compiled an initial listing of individuals with knowledge of genetic databases. We included people who serve as data stewards or as users of genetic testing data. We sought representatives in each of five key areas: health services research, clinical laboratory management, genetics, industry (e.g., development of genetic tests), and federal agencies. We also received suggestions of other candidates who met our criteria from individuals who could not participate in the interviews but wanted to make sure that their organization or their expertise in the genetic database arena was represented.

The goal of the interviews was to identify databases and mechanisms that might offer information about use of gene-based tests and interventions in healthcare in this country. Interviews, conducted by senior researchers at the National Association of Health Data Organizations (NAHDO) in Salt Lake City, Utah, consisted of mostly open-ended questions on predetermined topics, tailored to the type of individual that was being interviewed. (Appendix C presents the semistructured interview protocols.)

Topics for discussion included:

- Databases containing information on genetic tests and the scope of those data;

- Nature of variables in the databases;

- Level of access for research purposes;

- Feasibility of developing a surveillance system based on genetic tests; and

- Barriers, limitations, and advantages of such surveillance systems.

Another purpose of collecting information on existing databases from key informants was to identify critical stakeholders. These groups are ones that can provide additional inputs to the final design of any initiatives to establish a surveillance system and those that might be targeted for funding when and if an initiative is launched.

NAHDO completed 13 key informant interviews during July 2007 (see Appendix B for names and affiliations).

Laboratory Database Review

We also sought to understand the environment of genetic laboratories and their attendant databases, computer infrastructures, and related software programs. For this part of the assessment, we focused on molecular diagnostic laboratories to identify individuals who manage genetic testing data as part of their daily operations. These experts can be considered early adopters of databases for diagnostic data management and laboratory technologies including genetic testing. Our initial strategy was to identify laboratories that provide an assortment of genetic tests and that manage a significant amount of genetic test data. We searched for the websites of testing laboratories in March and April 2007 to identify these laboratories and find contact information for laboratory managers and directors who could describe their database and software resources. We found the major laboratories doing genetic tests through Internet searches using Google; the most successful phrase for searching was "molecular diagnostic laboratory." The next step was to identify key people who had knowledge of laboratory computer infrastructure and strategic decisionmaking process, such as the laboratory directors. Appendix B on key informants lists the individuals with whom we spoke specifically about laboratory databases.

Invitational Workshop

RTI organized the workshop, convened October 3-4, 2007, at the AHRQ Conference Center in Rockville, Maryland. The purpose was to elicit from participants guidance or possible recommendations for linking or modifying existing database systems or developing a new data system for monitoring and other uses set out by AHRQ. We worked closely with AHRQ staff, CDC liaisons, and TEP members to establish the goals of the meeting, a working agenda and topics for presentations and discussions, lists of possible participants, and other details and logistics.

Seven months before the workshop, we developed a provisional agenda for the 1.5-day meeting. The final program, which included both plenary and breakout working sessions, appears in Appendix D. The three breakout sessions covered:

1. Test utilization and knowledge of tests results

2. Impact on treatment choices and short-term clinical outcomes

3. Health and long-term outcomes.

Given the small size of the workshop (constrained by time and available resources), participant selection was a critical activity. Workshop participants, numbering approximately 30 individuals, included the following (apart from RTI project staff): TEP members; representatives of key stakeholder groups for this topic; additional experts, including our clinical consultant, to provide expertise on specific topics; and key staff from AHRQ, CDC, and other interested federal agencies, such as the National Cancer Institute (NCI), Food and Drug Administration (FDA), Centers for Medicare and Medicaid (CMS), and Health Resources and Services Resource Administration (HRSA). Appendix E provides the final list of participants.

We distributed a workshop book about 3 weeks before the meeting. It contained:

- Agenda and program;

- List of participants;

- A preliminary draft of this report (i.e., findings of our data collection activities and preliminary conclusions); and

- Brief "case studies" to help guide the three main breakout sessions noted above.

Workshop participants were asked to review the initial findings in the preliminary draft report, identify strengths and weaknesses, discuss various possibilities for monitoring gene-based interventions, and consider opportunities that may be presented by developments in health information technologies, such as electronic medical records and standards for linking and sharing electronic healthcare data. They were also asked to propose a research agenda that would guide development of such systems and allow ongoing assessment of the use and outcomes of gene-based clinical interventions in the US healthcare system.

Much of the substantive work of the workshop was conducted in the three separate workgroups. In order to focus the discussion at the workshop, a set of questions were provided to the participants in a case study format. The selected case studies covered a range of issues: prevalence (common or rare), type of genetic variation (inherited or acquired), and clinical context (treatment or prevention), including two distinct examples of pharmacogenomics. The examples were designed to focus the discussion on availability of information in the database on the following issues: (1) Patient population: asymptomatic persons or those diagnosed with a disease, (2) Test: type and results, and its utilization, (3) Impact on choice of intervention: screening/surveillance with another test such as mammography, surgery, or therapeutics, and (4) Impact on patient outcomes: mortality, morbidity. The selected case studies were intended to help identify to what extent information may be available from existing databases and to what extent we need to modify or create new databases and analytic tools. The selected case studies covered the following scenarios:

1. Common Acquired Genetic Variation/Prevention

2. Common Acquired Genetic Variation/Treatment

3. Common Inherited Genetic Variation/Prevention

4. Common Inherited Genetic Variation/Treatment

5. Rare Acquired Genetic Variation/Treatment

6. Rare Inherited Genetic Variation/Prevention

7. Pharmacogenomics.

Each workgroup was given a set of four case studies:

Workgroup 1: Test utilization and knowledge of tests results

1. **Common Acquired Genetic Variation/Prevention: PreGen Plus Test**
   Colorectal cancer screening: In asymptomatic persons, PreGen Plus test (detects APC, K-ras, p53 gene mutations) could be used for early detection of colorectal cancer. Information needed from the database includes: utilization of test; knowledge of test results.

2. **Common Inherited Genetic Variation/Prevention: *BRCA* Gene Mutations**
   Breast cancer: The *BRCA*1 and *BRCA*2 gene mutations in asymptomatic women with a family history of breast and/or ovarian cancer can predict high risk of breast cancer. Database information needed on utilization of tests and test results.

3. **Common Inherited Genetic Variation/Treatment: BRCA Mutations**
   Needed information from database: utilization of BRCA1 and BRCA2 testing *in women diagnosed with breast cancer* and knowledge of test results.

4. **Rare Inherited Genetic Variation/Prevention: PKU Screening**
   Phenylketonuria (PKU): Screening prenatally with DNA tests and of newborns for PKU to prevent mental retardation and other adverse health outcomes. Information needed from the database includes: type of test (prenatal DNA, newborn screening), utilization of tests and test results, including numbers of false-positives/true-positives.

Workgroup 2: Impact on treatment choices and short-term clinical outcomes

1. **Common Acquired Genetic Variation/Treatment: Gene Expression Profiling in Breast Cancer**
   Breast cancer: In women with breast cancer, MammaPrint and Oncotype Dx tests can potentially be used to clarify risk of recurrence of breast cancer. Information needed from the database includes: utilization of test and test results; how the test results influence downstream treatments such as additional chemotherapy and hormone

therapy; and the outcomes of therapy (impact on breast cancer mortality, quality of life).

2. **Rare Acquired Genetic Variation/Treatment: *BCR-ABL* Gene for CML**
   Chronic Myelogenous Leukemia (CML): *BCR-ABL* gene is a hallmark of CML. Additionally, the drug imatinib (Gleevec) is an inhibitor of the BCR-ABL protein that has become a first-line agent to treat early stage CML. Information needed from the database includes: utilization of *BCR-ABL* tests, outcomes of imatinib therapy (adverse events, CML-specific mortality).

3. **Pharmacogenomics – HER2/Neu Testing**
   Breast cancer: HER2 status is useful in determining treatment decisions including predicting for either resistance or sensitivity to different types of chemotherapeutic agents. Breast cancer patients who are HER2 positive have better prognosis and respond better to tyrosine kinase inhibitors than HER2 negative patients. Information needed from the database includes: patient populations (i.e., diagnosed with breast cancer); results of genetic testing; treatment decisions (e.g., use of tyrosine kinase inhibitors); outcomes.

4. **Pharmacogenomics -- Warfarin/Genetic Testing CYP2C9 and VKORC1 Genes**
   Warfarin/Genetic testing: Variations in CYP2C9 and VKORC1 genes can affect the dosing of warfarin. Information needed from the database includes: patient populations (i.e. diagnosed with DVT, PE or atrial fibrillation); results of genetic testing; final therapeutic dose of warfarin; INR levels; outcomes (incidence of strokes, thrombo-embolic and hemorrhagic, and downstream consequences, such as other bleeding events).

Workgroup 3: Health and long-term outcomes

1. **Common Acquired Genetic Variation/Prevention: PreGen Plus Test**
   Colorectal cancer screening: In asymptomatic persons, PreGen Plus test could be used for early detection of colorectal cancer. Information needed from the database includes: clinical interventions (colonoscopy, biopsies, treatment for colorectal cancer etc.) resulting from testing and impact on outcomes (incidence of colorectal cancer, colorectal cancer-mortality, and morbidity).

2. **Common Inherited Genetic Variation/Prevention: *BRCA* Gene Mutations**
   Breast cancer: The *BRCA*1 and *BRCA*2 gene mutations in asymptomatic women with a family history of breast and/or ovarian cancer can predict high risk of breast cancer. Database information needed on downstream interventions resulting from testing: counseling, prophylactic surgery (mastectomy or oophorectomy), frequent surveillance (MRI, mammography etc.) or prophylactic drugs (tamoxifen). Information is also needed on outcomes of these interventions such as morbidity and adverse effects of surgery, drugs or surveillance; incidence of breast cancer; and breast cancer mortality.

3. **Common Inherited Genetic Variation/Treatment: BRCA Mutations**
   Breast cancer: Information from database needed to answer whether natural history of breast cancer is different (more aggressive, more resistant) in women with *BRCA*1 and *BRCA*2 gene mutations compared to those without mutation. Additionally, information is needed to understand if treatment were different in women with these mutations and the outcomes of these treatments.

4. **Rare Inherited Genetic Variation/Prevention: PKU Screening**
   Phenylketonuria (PKU): Screening prenatally with DNA tests and of newborns for PKU to prevent mental retardation and other adverse health outcomes. Information needed from the database includes: compliance of patients with dietary restrictions and long-term outcomes, including incidence of mental retardation and quality of life of patients and family members.

Using the case studies as an initial basis for discussion, workgroups were given the following charge:

- Determine to what extent information may be available in current databases to address specific questions (i.e., as put forward in the case studies).

- Develop options for using or modifying existing databases for specific purposes.

- Summarize key findings that cut across the different case study scenarios.

- Develop recommendations for additional research and development.

A designated rapporteur from each workgroup summarized the findings and presented them for review and comment by all workshop participants at a plenary session the following day.

Peer Review

After the workshop, the RTI DEcIDE team revised and completed the draft report to take account of all the workshop deliberations, conclusions, and recommendations. The report was then subjected to formal, external peer review by about 18 outside experts in the field (who were not participants of the workshop). Potential peer reviewers were contacted 2 months ahead of time to ensure that they could accommodate our schedule and to acknowledge that they are agreeing to be listed in the report as peer reviewers. We asked for reviews to be returned within 4 weeks. As comments were received, we entered them into a database and recorded the RTI team's responses or changes as part of the formal Peer Review Disposition Report submitted to AHRQ. We then produced the final report for submission to AHRQ and placement on the AHRQ website.

Appendix B. Key Informants

Key Informant Interviewees

Sarah Carr, PhD
Office of the Director
National Institutes of Health:
Associate Director for Policy
Clinical Research Policy Analysis and Coordination
Bethesda, MD

Cathy Fomous, PhD
Senior Health Science Policy Analyst
National Library of Medicine
Bethesda, MD

Ms Suzanne Goodwin
Office of Biotechnology Activities
Secretary's Advisory Committee on Genetics, Health, and Society
NIH Office of Biotechnology Activities
Bethesda, MD

Robert L. Davis, MD
The Center for Health Research - Southeast
Kaiser Permanente
Atlanta, GA

Felix W. Frueh, PhD
Associate Director for Genomics
Office of Clinical Pharmacology
Center for Drug Evaluation and Research/Food and Drug Administration
Silver Spring, MD

Muin J. Khoury, MD, PhD
Director
National Office of Public Health Genomics
Coordinating Center for Health Promotion
Centers for Disease Control and Prevention
Atlanta, GA

Catherine McCarty, PhD, MPH
Senior Research Scientist and Interim Director
Center for Human Genetics
Marshfield Clinic Research Foundation
Marshfield, WI

Richard Platt, MD, MS
Harvard Pilgrim Health Care
Harvard Medical School and
Brigham and Women's Hospital
Boston, MA

David Veenstra, PharmD, PhD
Pharmaceutical Outcomes Research and Policy Program
Department of Pharmacy,
Institute for Public Health Genetics
University of Washington
Seattle, WA

Michael S. Watson, PhD, FACMG
Executive Director
American College of Medical Genetics
Bethesda, MD

Louise Wideroff, PhD, MSPH
Risk Factor Monitoring and Methods Branch
Applied Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute
Bethesda, MD

Andrew M. Wiesenthal, MD, SM
Associate Executive Director
The Permanente Federation
Oakland, CA

Marc S. Williams, MD
Director
Intermountain Clinical Genetics Institute
Salt Lake City, UT

Raymond Woosley, MD, PharmD
President and Chief Executive Officer
The Critical Path Institute
Tucson, AZ

Level 2 Laboratory Database Interviewees

Alexis B. Carter, MD
Director of Pathology Informatics
Emory University School of Medicine
Atlanta, GA

Dan Roden, MD
Vanderbilt University School of Medicine
Nashville, TN

Dan Jones, MD, PhD
Director
Molecular Diagnostics Lab
MD Anderson Cancer Center
Houston, TX

Joyce A. Mitchell, PhD, FACMI, FACMG
Chair, Dept of Medical Informatics
University of Utah
Salt Lake City, UT

Jonathan F. Tait, MD, PhD
Co-Director of Genetics
University of Washington
Seattle, WA

Andrea Ferreira-Gonzalez, PhD
Secretary's Advisory Committee on Genetics, Health
and Society
Virginia Commonwealth University
Richmond, VA

Jeffrey A. Kant, MD, PhD
University of Pittsburgh Medical Center
Pittsburgh, PA

Appendix C. Key Informants Discussion Guidelines-Protocol

The script: For database "owners" who developed or maintain database(s)

Hello Mr./Dr./Ms. _____; (proper salutations and greetings)

I am calling you on behalf of National Association of Health Data Organizations also known as NAHDO, in regards to a project commissioned by the Centers for Disease Control and Prevention (CDC) and the Agency for Healthcare Research and Quality (AHRQ). The purpose of the project is to identify databases and mechanisms that might offer information about the utilization of gene-based tests and interventions in healthcare in this country. We will use this information to develop guidance, research priorities, and recommendations for establishing a monitoring system(s) for gene-based tests and interventions in the U.S. healthcare system. We are contacting you because of your knowledge of one or more databases about genetic testing.

Key Issues:

- Identify Data Bases
- Currency of database(s)
- Nature of interviewee's involvement: (primary responsibility or other role)
- The primary purpose(s) of the database
  - Specify tests included in the database: _____
- Data availability for other uses, such as research or public health surveillance
  - Any restrictions for its use
- Level and scope of data: e.g. covers patients of the healthcare organization, state, regional, national, or other identifiable population
- Variables in the dataset and the possibility of linkage with intervention and outcomes
- Technical characteristics of data (use of standard terms, billing codes, medical controlled vocabulary)
- Possibility of sharing these data with a federal agency
- Contact person/agency responsible for maintaining this database
- Potential of gene-based test database for monitoring test utilization and related interventions in health care in this country
- Is it desirable to establish a monitoring system for gene-based tests and interventions?
  - Benefits, Scale,
  - Reasons, if not desirable
- Knowledge of existing data systems of genetic tests
  - Sufficient for establishing a monitoring system(s) for gene-based tests and interventions, at local, state, or national level?
- If they are not sufficient: What future data developments are needed to create and maintain a monitoring system(s) for gene-based tests and interventions?
- Technical and non-technical (legal) issues in developing an effective monitoring system for gene-based tests and interventions.

The script: For user of databases for public health, health services research, or policy

Hello Mr./Dr./Ms. _____; (proper salutations and greetings)

I am calling you on behalf of National Association of Health Data Organizations also known as NAHDO, in regards to a project commissioned by the Centers for Disease Control and Prevention (CDC) and the Agency for Healthcare Research and Quality (AHRQ). The purpose of the project is to identify databases and mechanisms that might offer information about the utilization of gene-based tests and interventions in healthcare in this country. We will use this information to develop guidance, research priorities, and recommendations for establishing a monitoring system(s) for gene-based tests and interventions in the U.S. healthcare system. We are contacting you because of your knowledge of one or more databases about genetic testing.

Key Issues:

- Identify Data Bases
- Currency of database(s)
- Nature of interviewee's involvement: (primary responsibility or other role)
- The primary purpose(s) of the database
    - Specify tests included in the database: _____
- Data availability for other uses, such as research or public health surveillance
    - Any restrictions for its use
- Restrictions on use of this genetic testing database
- Level and scope of data: e.g. covers patients of the healthcare organization, state, regional, national, or other identifiable population
- Variables in the dataset and the possibility of linkage with intervention and outcomes
- Technical characteristics of data (use of standard terms, billing codes, medical controlled vocabulary)
- Possibility of sharing with a federal agency
- Contact person/agency responsible for maintaining this database
- Potential of gene-based test database for monitoring test utilization and related interventions in health care in this country
- Is it desirable to establish a monitoring system for gene-based tests and interventions?
    - Benefits, Scale,
    - Reasons, if not desirable
- Knowledge of existing data systems of genetic tests
    - Sufficient for establishing a monitoring system(s) for gene-based tests and interventions, at local, state, or national level?
- If they are not sufficient: What future data developments are needed to create and maintain a monitoring system(s) for gene-based tests and interventions?
- Technical and non-technical (legal) issues in developing an effective monitoring system for gene-based tests and interventions.

Appendix D. Draft Workshop Agenda

Establishing an Infrastructure for Monitoring Use and Outcomes of Gene-Based Therapies

October 3-4, 2007
John M. Eisenberg Building
Agency for Healthcare Research and Quality Conference Center
Rockville, Maryland
Center for Outcomes and Evidence (COE) Conference Room

<u>**Wednesday, October 3, 2007 (AHRQ Conference Room)**</u>

7:30 a.m.   Transportation from Hotel to Eisenberg Building

8:00 a.m.   Continental Breakfast Available

8:30 a.m.   **Welcome, Introductions, Housekeeping, Goals for the Day**
                       Frank DeStefano, MD, MPH
                       Scott Smith, PhD or Gurvaneet Randhawa, MD, MPH

9:00 a.m.   **Background: Issues Driving this Project from the AHRQ and CDC Perspectives**
                             Scott Smith, PhD or Gurvaneet Randhawa, MD, MPH
                             Ralph Coates, PhD

9:20 a.m.   **RTI DEcIDE Draft Report: Overview and Audience Discussion**
                             Frank DeStefano, MD, MPH

10:10 a.m.   **Charge to the Working Groups**
                             Frank DeStefano, MD, MPH
                                1. *Brighton Dam and Watts Branch Conference Rooms*
                                2. *Great Fall Conference Room*
                                3. *Teleconference Room 1101*
10:15 a.m.   Break

10:30 a.m. – 3:30 p.m. **Breakout Sessions** *(moderated working groups)*
                             Background and Objectives
                             Discuss Case Studies
                             Lessons from Case Studies
                             Preliminary Conclusions and Recommendations

12:00 – 12:30 p.m. Working lunch in the breakout rooms:

                       Workgroup 1: *Options and Approaches for Monitoring Utilization of Gene-Based Tests*
                             Great Fall Conference Room
                             Moderator: Watson
                             Rapporteur: Whitehead

Workgroup 2: *Options and Approaches for Monitoring Trends of Gene-Based Treatments (Pharmacogenomics)*
>> **Teleconference Room 1101**
>> *Moderator: Murff*
>> *Rapporteur: Masica*

Workgroup 3: *Options and Approaches for Monitoring Outcomes of Gene-Based Applications*
>> **Brighton Dam and Watts Branch Conference Room**
>> *Moderator: McCarty*
>> *Rapporteur: DeStefano*

3:30 p.m.	Adjourn for day (dinner on one's own)
3:45 p.m. - as needed*:*	Workgroup moderators and rapporteurs: Prepare workgroup summary and presentation for next day

Thursday, October 4, 2007 (AHRQ Conference Room)

8:00 a.m.	Transportation from Hotel to Eisenberg Building (check out of hotel)
8:30 a.m.	Continental Breakfast Available
9:00 a.m.	**Welcome for Day 2; Housekeeping; Goals for the Day** *Frank DeStefano*
9:15 – 11:45 a.m.	**Reports from the Workgroups: Provisional Findings and Conclusions** *Moderator: Frank DeStefano*
9:15 a.m.	**Workgroup 1: Monitoring Utilization of Gene-Based Tests: Presentation and Audience Discussion** *Rapporteur: Nedra Whitehead*
10:00 Break	
10:15 a.m.	**Workgroup 2: Monitoring Trends of Gene-Based Treatments (Pharmacogenomics): Presentation and Audience Discussion** *Rapporteur: Andrew Masica*
11:00 a.m.	**Workgroup 3: Monitoring Outcomes: Presentation and Audience Discussion** *Rapporteur: Frank DeStefano*
11:45 a.m.	**Plenary discussion:** *Guidance and Recommendations* **Provisional Recommendations Research Agenda** *Moderator: Frank DeStefano*
12:15 p.m.	**Wrap-up Comments**
12:30 p.m.	Symposium Adjourn

Appendix E. Workshop Participants

David Adamson, MD, FRCSC, FACOG, FACS
Fertility Physician
Arcfertility
Fertility Physicians of Northern California
Palo Alto, CA 94301

Christopher Chute, MD, DPH
Chair
Biomedical Informatics
Mayo Clinic College of Medicine
Rochester, MN 55905

Ralph J. Coates, PhD
Acting Associate Director for Science
National Office of Public Health Genomics
NCCDPHP, CCHP
Centers for Disease Control and Prevention
Atlanta, GA 30341-5535

Beryl Crossley, BM, BCh
Computational Biologist
Genetics R&D
Quest Diagnostics Nichols Institute
San Juan Capistrano, CA 92675

Maryellen de Mars, PhD
Director
Clinical Biomarkers
The Critical Path Institute
Rockville, MD 20850

Anne Elixhauser, PhD
Health Services Research
Center for Delivery, Organization, and Markets
(CDOM)
Agency for Healthcare Research and Quality (AHRQ)
Rockville, MD 20852

Paul Fernhoff, MD, FAAP, FACMG
Geneticist
Emory Genetics
Decatur, GA 30033

Felix Frueh, PhD
Associate Director for Genomics
Office of Clinical Pharmacology, CDER/FDA
Rockville, MD 20857

Kathy Helzlsouer, MD, MHS
Director, Prevention and Research Center
Weinberg Center For Women's Health & Medicine
Mercy Medical Center
Baltimore, MD 21202

Mark Hornbrook, PhD
Chief Scientist
Center for Health Research
Kaiser Permanente
Northwest Region
Portland, OR 97227

Louis Jacques, MD
Director, Division of Items and Devices
Coverage & Analysis Group, OCSQ
Centers for Medicare & Medicaid Services
Baltimore, MD 21244

Michele Lloyd-Puryear, MD, PhD
Chief Genetics Services Branch
Maternal and Child Health Bureau,
Health Resources and Services Administration
Rockville, MD 20857

Deborah Loer-Martin, PhD, CRA
Director of Grants & Research
Health Choice Network
Pediatric Asthma Program
Miami, FL 33166

Rochelle Long, PhD
Chief, Pharmacological & Physiological Sciences
Branch
Pharmacology, Physiology, & Biological Chemistry
Division
National Institute of General Medical Sciences
Bethesda, MD 20892-6200

Andrew Masica, MD, MSc
Clinical Scholar/Internal Medicine-Hospitalist
Baylor Institute for Health Care Research and
Improvement
Baylor Health Care System
Dallas, TX 75206

Catherine McCarty, PhD, MPH
Senior Research Scientist and Interim Director,
The Center for Human Genetics
Marshfield Clinic Research Foundation
Marshfield, WI 54449

Harvey Murff, MD, MPH
Research Safety Advocate
Division of General Internal Medicine and Public
Health
Institute of Medicine and Public Health Sixth Floor,
Suite 600 1738
Vanderbilt University
Nashville, TN

Kathryn Phillips, PhD, MPA
Member of UCSF Comprehensive Cancer Center
Professor of Health Services Research and Health
Economics
School of Pharmacy
University of California, San Francisco
San Francisco, CA 94143

Margaret Piper, PhD, MPH
Associate Director
Blue Cross Blue Shield Association Technology
Evaluation Center
Atlanta, GA 30306

Parminder Raina, PhD
Evidence Based Practice Centre Director
McMaster University
Ontario, Canada L8S 4L8

Gurvaneet Randhawa, MD, MPH
Task Order Officer
Senior Advisor on Clinical Genomics
Center for Outcomes and Evidence (COE)
Agency for Healthcare Research and Quality
Rockville, MD 20850

Maren Scheuner, MD, MPH
Natural Scientist
Senior Research Scientist, UCLA Center for Health
Policy Research
Visiting Associate Professor, UCLA School of Public
Health
RAND Corp
Santa Monica, CA 90407-2138

Scott Smith, PhD, MPH
Project Officer for DEcIDE Program
Center for Outcomes and Evidence (COE)
Agency for Healthcare Research and Quality
Rockville, MD 20850

R. Grant Steen, PhD
President
Medical Communications Consultants, LLC
Chapel Hill, NC 27517

Michael Watson, PhD, FACMG
Executive Director
American College of Medical Genetics (ACMG)
Bethesda, MD 20814-3998

Louise Wideroff, PhD, MSPH
Program Director
Risk Factor Monitoring and Methods Branch
Applied Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute
Bethesda, MD 20892-7344

Ann Willey, PhD, JD
Director
Office of Laboratory Policy and Planning
Wadsworth Center, NYSDOH
Albany, NY 12223

Athanasios Zavras, DMD, MS, DrMS
Molecular Epidemiologist
HSDM Associate Professor of Health Policy &
Epidemiology
Harvard Faculty of Medicine - School of Dental
Medicine
Boston, MA 02115

RTI International

Amali Amarasinghe, MPH
Clinical Studies Specialist
RTI International
Washington, DC 20005-3967

Frank DeStefano, MD, MPH
Senior Research Epidemiologist
RTI International
Atlanta, GA 30341-5533

Jo Hetland
Project Coordinator
RTI International
Atlanta, GA 30341-5533

Chris McClure, MPH
Research Epidemiologist
RTI International
Rockville, MD 20852

Nedra Whitehead, PhD, MS
Senior Genetic Epidemiologist
RTI International
Atlanta, GA 30341

Appendix F. Literature Search Abstraction Forms

Article Information	
ID:	Reference Manager Database ID
Authors:	Include all authors; full names if provided
Title:	Article title
Citation	Journal name. volume #(Issue #):pages
Publication Date:	Month Year
PubMed ID	From abstract

Organization Information	
Type of Facility	Laboratory, clinic, hospital, public health agency
Private/Public	Private entity such as corporation or private university versus government owned
Contact Info	
Years of experience	Years of experience performing test
Number of tests done	Total number of test performed within the reported time

General Database Information	
Publicly Available?	Is the database available to external analysts through public access datasets, web accessed query systems, or data sharing agreements?
Proprietary?	Does the owner of the database consider any information within the data base proprietary?
Types of Data	Categories of data, for example: demographic data, test results,
Potential Public Health Use of Data	List ways the data may be useful to public health agencies in monitoring use and outcomes of gene-based tests or in policy or health services research.
Population	Are the data drawn from an identifiable population? If so, briefly describe the population.
Notes:	

Database Details			
Variables	*Coding*	*Sources*	*Useful for Public Health?*
List all variables available in the article	List any information on the coding of the variables in the database.	List the source of the data included in the database	

Potential Linkages		
Database	*Added information*	*Linking variables*
Databases that it may be possible to link to	What additional information would be available from these databases? Medical interventions or outcomes Public health measures (ie, prevalence; incidence) Population denominator data	Variables through which data would be linked: Individual identifiers Geographic units Clinical or research ID numbers

Article Information	
ID	Reference Manager Database ID
Authors	Include all authors; full names if provided
Title	Article title
Citation	Journal name. volume #(Issue #):pages
Publication Date:	Month Year
PubMed ID	From abstract

Organization Information	
Type of Facility	Laboratory, clinic, hospital, public health agency
Private/Public	Private entity such as corporation or private university versus government owned
Contact Info:	
Years of experience	Years of experience performing test
Number of tests done	Total number of test performed within the reported time
Estimated % of market	The percentage of the total number of given test(s) performed in a given population that are performed by this organization. Serves as estimated population coverage for this data
How estimated	Description of how above number was derived.

Information on Test(s)	
Condition	The condition or conditions diagnosed or screened for in the given tests
Screening/Diagnostic	Both
Total Tests Performed	Total number of tests performed as reported in the article
Total Time Period	Time covered by article
Time Trends	Any available information on changes in test utilization over time. In order of preference: 1. Distribution over time; 2. Description of changes

Utilization Data Details May need to be modified depending on test, condition, and information provided in the article

Postnatal Testing If relevant, present postnatal and prenatal testing data separately.

Demographic Variables	Number	%		Test Results	Notes
Sex					
Male					
Female					
Age					
Race					
Prenatal Testing					
Male					
Female					
Age					
Race					

Appendix G. Additional Databases

The following databases were suggested by peer reviewers of the report as containing potentially relevant information on gene-based tests or other gene-based applications:

Collaborative Education and Test Translation (CETT) program

National Institutes of Health (NIH)-supported pilot program to compile clinical and genetic data for individuals with selected rare diseases. Its primary purpose is to identify genotype-phenotype associations.

DuchenneConnect

A web-based service to serve as a central hub linking the resources and needs of the Duchenne/Becker muscular dystrophy community: those living with the disease; family, friends, and caregivers; and the medical research community. Is developing a patient registry.

Genetic Alliance BioBank

Advocacy group biobank that provides a state-of-the-art storage facility and system for the collection and archiving of DNA, tissue, and cell lines. Includes an informatics core that encodes identifiers in a centralized database. The biobank and database are owned and maintained by each advocacy organization.

Genzyme

A large commercial laboratory of reproductive genetic tests.

MarketScan Databases

A proprietary U.S. research database, the MarketScan® data warehouse contains fully integrated, de-identified, individual-level healthcare claims data (inpatient, outpatient, drug, lab, health risk assessment, and benefit design) from commercial, Medicare supplemental, and Medicaid populations.

Medicaid

A state administered health care coverage program available to certain low-income individuals and families. Some states have Medicaid beneficiaries enrolled in integrated healthcare systems with electronic health records. Medicaid claims and other data, in at least some states, may be accessible for public health or health services research.

Muscular Dystrophy Surveillance Tracking and Research Network (MD STARnet)

The Centers for Disease Control and Prevention (CDC) is working with researchers in Arizona, Colorado, Georgia, Iowa, and western New York to set up surveillance/tracking systems for Duchenne/Becker muscular dystrophy (DBMD). The goal of the project is to find all DBMD patients in these states by using information from different sources, such as clinic medical records and hospital records.

TREAT-NMD patient registries

A pan-European and U.S. collaborative of databases/registries for Duchenne muscular dystrophy, spinal muscular atrophy, and other muscular dystrophies. Patients will be registered with their genetic defects and clinical status. The primary objective of the TREAT-NMD registries is the facilitation, planning and recruitment of clinical trials.

Tuberous Sclerosis Complex (TSC) Natural History Database

A voluntary patient registry with the goal of helping to understand the natural history of the disease. A collaborative project of the Tuberous Sclerosis Alliance and a consortium of TSC clinics.

**U.S. Department of
Health and Human Services**

Public Health Service
Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, MD 20850