The Value of Laboratory Screening and Diagnostic Tests for Prevention and Health Care Improvement

Prepared for:

American Clinical Laboratory Association and
Advanced Medical Technology Association (AdvaMed)

Prepared by:

The Lewin Group, Inc.

September 2009

The Value of Laboratory Screening and Diagnostic Tests for Prevention and Health Care Improvement

Prepared for:

American Clinical Laboratory Association and Advanced Medical Technology Association (AdvaMed)

Prepared by:

The Lewin Group, Inc.

September 2009

Table of Contents

Overview

The contributions of clinical laboratory screening and diagnostic tests[a] to health care quality and outcomes are substantial. These contributions were described in an earlier report from The Lewin Group, *The Value of Diagnostics Innovation, Adoption, and Diffusion in Health Care* (2005). This report updates key elements of that study, providing a current overview of the important role of laboratory screening and diagnostic tests in our health care system, today's means of assessing value, and four case studies documenting value of specific tests to patient care.

Our overarching finding is that: *Innovation, demonstrated clinical benefit, and appropriate use of laboratory screening and diagnostic tests are essential for achieving the goals of health system reform. Clinical laboratory testing is integral to evidence-based improvements in health care quality, patient outcomes, efficiency, and accountability*. This finding is substantiated, in part, by the four case studies accompanying this report:

- Rapid methicillin-resistant *Staphylococcus aureus* (MRSA) testing for identifying health care-acquired infections

- Hemoglobin A1c (HbA1c) testing for screening and diagnosis of prediabetes and diabetes

- *KRAS*[b] gene mutation testing for targeted treatment of colorectal cancer

- Human papillomavirus DNA (HPV) testing to screen and diagnose cervical cancer

Though each of these case studies represents a unique configuration of health problems, at-risk patient populations, and testing technology, some common capabilities of laboratory testing emerge. Examples of these common elements include ongoing innovation, evidence of clinical benefit, early detection and treatment to control disease, better targeting and accuracy that enable more effective management and decision-making, more efficient care, and cost-saving opportunities. These case studies underscore the role of laboratory medicine in augmenting the evidence base for health care, decision-support for clinicians and patients, prevention and wellness, better patient outcomes, and better value.

Many of the benefits of laboratory testing are not being realized in the current system. The main body of the report reviews the current processes for assessing the value of laboratory screening and diagnostic tests and key policy issues that limit the realization of optimal value. Among the hurdles to appropriate use of these services are: insufficient provider awareness regarding when to use tests, challenges in generating evidence regarding the clinical utility of tests for particular patient subgroups and indications, inconsistencies among clinical practice guidelines regarding appropriate use of tests, inconsistent or inadequate coverage and payment policies, and need to provide additional evidence of the favorable economic impact of laboratory testing.

---

[a] Depending on the patients' needs and the level of technical and resource capabilities required to perform the test, clinical laboratory screening and diagnostic tests may be conducted in laboratories (including independent reference laboratories, hospital and academic medical center laboratories, physician office laboratories, nursing home laboratories, and others) or at the point-of-care.

[b] *v-Ki-ras2 Kirsten rat sarcoma viral oncogene*

I.   Introduction

Although the U.S. leads the world in per capita and total health care spending and is the world's greatest source of health care innovation, it ranks poorly among industrialized nations in two gross indicators of health care quality: infant mortality and life expectancy.[1] As documented in recent landmark studies, only 55% of adults and 47% of children in 12 metropolitan areas across the U.S. received recommended care, contributing to the significant shortfalls in overall care quality, patient safety, and health outcomes.[2, 3] There are similar deficits in adherence to recommendations for screening and follow-up care—only 62% of recommended laboratory and radiology tests were provided for preventive, acute, and chronic care. Aside from their implications for patient health, these shortfalls contribute to the sizable unnecessary spending in health care in the U.S., amounting to as much as $0.30 or more of every health care dollar.[4] In the current health care environment, the needs and opportunities to improve health care quality and efficiency are great and immediate.

A.   Increasing Value is a National Priority

According to the Congressional Budget Office (CBO), continued escalation of health care costs is the main long-term threat to the federal budget and the nation's overall financial well-being.[5] Increasing value in the health system and reducing unnecessary spending has become a national priority.   Since publication of the Institute of Medicine (IOM) report, *Crossing the Quality Chasm: A New Health Care System for the 21st Century* (2000), stakeholders have been directing reform efforts toward strategies that increase the value of our health care services and produce better patient outcomes per dollar spent.[6] The intent of this approach is to align how value is created for patients across services and time, and to target medical conditions over the cycle of care (episodes of care) and differing levels of care. Health care system stakeholders (including payers, providers, patients, accreditation and quality improvement organizations, health services researchers, policymakers, industry) agree that increasing value to patients can be achieved by improving quality through the types of goals shown in Box 1.[7] Laboratory screening and diagnostic tests are essential for pursuing many of these goals.

Box 1
Health System Goals for Increasing Quality and Value

- Prevention
- Early detection
- Right diagnosis
- Early treatment
- Right treatment to the right patients
- Treatment earlier in causal chain of disease
- Fewer delays in the care delivery process

- Less invasive treatment methods
- Faster recovery
- More complete recovery
- Less disability
- Fewer relapses or acute episodes
- Slower disease progression
- Less need for long-term care[6]

B. Value of Laboratory Medicine to Clinical Care

The contributions of clinical laboratory screening and diagnostic tests to the health system goals (listed in Box 1) are substantial. As an essential component of high quality of care, laboratory tests are used for much more than diagnosis of disease in symptomatic individuals. Laboratory testing is an integral part of many medical decisions, providing clinicians with often pivotal information necessary for prevention, diagnosis, treatment, and management of disease.[8, 9] Despite the extensive role of laboratory medicine in informing medical decision-making, in 2007, spending on Part B laboratory services was $6.8 billion or just 1.6% of total Medicare expenditures[c] [10, 11] and 2.3% of national health care spending. [d] [11, 12] Significant contributions of laboratory medicine remain untapped.

The value of laboratory medicine is realized through its many roles in patient care. These include screening of asymptomatic individuals to identify risk for developing disease, detecting disease at the earliest stages before symptoms occur, selecting safe and effective treatments, planning disease management strategies, estimating treatment response throughout the course of care, identifying threats to patient safety and public health, such as hospital-acquired infections (HAI), protecting the blood supply and transplant recipients from harmful pathogens, and drugs of abuse testing to support clinical care and assure public safety. These aspects of value can be expressed along a continuum of care such as is listed in Box 2.[13]

> **Box 2**
> **Value of Laboratory Tests in Clinical Decision-making**
> - Screen for disease
> - Screen to determine risk for developing disease
> - 'Rule in' of a diagnosis
> - 'Rule out' of a diagnosis
> - Start an intervention
> - Adjust an intervention
> - Stop an intervention
> - Assess efficacy of an intervention
> - Assess compliance with an intervention
> - Assess prognosis

Laboratory medicine also is important to clinical guidelines. As described in our 2005 report, a search of clinical practice guidelines across 23 main condition/disease categories found that 37% focused on or involved laboratory tests.[14] Increasingly, the objective, scientific data produced by clinical laboratory tests is used to measure provider performance (individual and organizational) as well as to implement value-based purchasing that aims to optimize use of health care resources and decrease short-, medium-, and long-term costs of care.[15]

Additionally, the recent major advances in science and technology, such as those associated with molecular-level and genetic testing, are leading to changes in clinical practice. Genetic tests are now available for more than 1,700 diseases, up from about 1,250 in 2005.[16] New testing techniques tend to be more sensitive and specific, allowing clinicians to detect, diagnose, and manage disease more effectively. For example, technologies that rely on DNA, RNA, and protein composition allow evaluation of disease states at the molecular level, supporting earlier detection and a more personalized approach to patient care.[17]

---

[c] Total Medicare expenditures were $433 billion in 2007.[10]

[d] Total laboratory industry revenues were estimated at $51.7 billion in 2007; this is 2.3% of the $2.24 trillion in 2007 national health care expenditures estimated by the Centers for Medicare and Medicaid Services.[12]

Other innovative rapid testing techniques, such as those for microbiology, make testing for HAIs more cost-effective and less time-consuming, supporting real-time decision-making and operational efficiency.[18] Innovations in miniaturization have expanded the point-of-care testing (POCT) menu, enabling laboratory testing at the hospital bedside, physician's office, other clinical settings, and, in some cases, patient self-testing at home. These new technologies are demonstrating their value to improved patient outcomes and quality of life, fewer side effects of treatment, and decreased costs of care.

Examples of specific laboratory screening and diagnostic tests that can contribute to health system value are presented in Table 1.

Table 1. Examples of Screening/Diagnostic Tests and Services that Support the Health System Goals for Improving Value and Quality

Health System Goal*	Examples of Screening and Diagnostics Tests
Prevention of disease	• **HbA1c testing** to screen for prediabetes to identify at-risk patients and prompt interventions that can prevent onset of type 2 diabetes • **Factor V Leiden genetic variation** to identify increased risk of blood clots. • **Malignant hyperthermia genetic test to** identify individuals who will react to general anesthesia (see GeneTests for reference) • **Familial adenomatous polyposis genetic test to** identify individuals who will develop colon cancer and allow prophylactic colectomy or other surgical procedure
Right diagnosis	• **Cardiac enzyme marker tests,** e.g., for troponin, myoglobin, creatinine phosphokinase (CPK), which are released after a heart attack and identify heart damage • **Fragile X syndrome** test for a form of inherited mental retardation and developmental delay, to determine appropriate management and risk of familial recurrence.
Early detection and treatment	• *BRCA1* and *BRCA2* test to identify increased risk of breast cancer and ovarian cancer in the absence of tumor • **Prenatal and newborn screening** for inherited disorders, to enable initiation of treatment and to reduce adverse effects • (Intra-amniotic infection/inflammation**) IAI test**, a non-invasive cervical-vaginal fluid (CVF) proteomic biomarker test, to diagnose intra-amniotic infection in preterm labor, a condition that is 80-90% asymptomatic
Right treatment to the right patients	• **HER-2/***neu* (human epidermal growth factor receptor 2, also known as ErbB-2, *ERBB2*) **protein testing** for patients with breast cancer identifies those who will benefit from targeted treatment with trastuzumab (Herceptin) • *KRAS* **gene mutation testing** for patients with metastatic colorectal cancer distinguishes between patients who are most unlikely to benefit from, and those who are more likely to benefit from, the drug cetuximab • *BCR/ABL* **oncogene testing** for patients with chronic myelogenous leukemia who will benefit from treatment with imatinib (Gleevec)
Fewer mistakes and repeats in treatment	• **HIV viral load test** to determine disease progression and whether the drug is working • **Emphysema gene test** to identify likelihood of liver disease in emphysema patients without biopsy and allows early intervention
Fewer delays in the care delivery process	• **Rapid molecular MRSA testing** for *mecA* gene markers to identify, within two hours, patients with antibiotic-resistant *S. aureus* infections to guide drug selection and timely implementation of control measures • **Point-of-care tests** • **Developmental delay/mental retardation** tests, such as chromosome analysis, array comparative genomic hybridization (aCGH), biochemical genetic tests to help determine cause and appropriate therapies

Health System Goal*	Examples of Screening and Diagnostics Tests
Less invasive treatment methods	• **Gene expression profiling of RNA** isolated from peripheral blood mononuclear cells for noninvasive identification of heart transplant recipients with low probability of organ rejection • **Hereditary hemochromatosis** gene test to identify individuals who have the disease and decrease need for liver biopsy • **Cytochrome 2C19 gene test for clopidogrel** to identify heart attack or stroke patients who may not respond to therapy so that other therapies can be prescribed
Faster and more complete recovery	• **Breast cancer gene marker assays** to assess the aggressiveness of the tumor, help tailor treatment and predict risk of recurrence • **Hereditary hemochromatosis** gene test to identify individuals at risk for excess iron absorption and potential accompanying morbidity; test can replace liver biopsies in some patients
Less disability	• **Estimated glomerular filtration rate** for early detection of chronic kidney disease and monitoring of kidney function, reducing disability • **Lead screening** for early detection of elevated blood levels in children reduces risk of central nervous system damage
Fewer relapses or acute episodes	• **Sickle cell anemia** testing to reduce risk of sickle cell anemia crisis by early detection • **Cytochrome P450** testing to identify individual rate of drug metabolism thus reducing risk of adverse drug reactions • **International Normalized Ratio (INR)/prothrombin time testing** to monitor patients who are on warfarin therapy to prevent blood clots or severe internal bleeding
Slower disease progression	• **C-reactive protein testing** to detect individuals at high risk for serious cardiovascular events and candidates for statin therapy • **Prostate specific antigen testing** to screen at-risk men and monitor increases in year-to-year PSA values, which can identify prostate cancer that needs specific clinical intervention as opposed to watchful waiting • **Cytochrome P450 2D6 testing** to identify breast cancer patients who may not respond to tamoxifen therapy so that other therapies such as aromatase inhibitors can be prescribed • **HIV genotyping** to determine which antiviral drug combination is best
Less need for long-term care	• **Antinuclear antibody test** to detect lupus erythematosus, rheumatoid arthritis and other autoimmune disorders and help to determine appropriately targeted therapy • **Maternal serum screening** to identify pregnancies that could benefit from prenatal testing for spina bifida, Down syndrome and certain other birth defects

*Health system goals adapted from Porter (2007)

Sources
1. Marchant J. Innovations in diagnostics: next generation molecular and point-of-care diagnostics driving personalized healthcare. London, England: Business Insights Limited, 2005.
2. Business strategies for molecular diagnostics in the lab, including state of the market report 2007. Terry M, Murg S, ed. New York, NY: Washington G-2 Reports, 2007.
3. A purchaser's guide to clinical preventive services: moving science into coverage. National Business Group on Health and Centers for Disease Control and Prevention, 2009.
4. GeneTests. National Center for Biotechnology Information, National Institutes of Health, 2009. (Accessed July 20, 2009, at http://www.ncbi.nlm.nih.gov/sites/GeneTests/.)
5. LabTests Online. American Association of Clinical Chemistry, 2009. (Accessed July 20, 2009, at http://www.labtestsonline.org/.)
6. National Guideline Clearinghouse. Agency for Healthcare Research and Quality, 2009. (Accessed July 20, 2009, at http://www.guideline.gov/.)
7. Terry M. Lab industry strategic outlook: market trends and analysis 2007. New York, NY: Washington G-2 Reports, 2007.

C. Differentiating Screening and Diagnostic Tests

Clinical laboratory tests serve two overarching functions—screening and diagnosis:

- Laboratory tests used for **screening** assess the likelihood of the presence of a disease or condition in apparently healthy or **asymptomatic** individuals who are at sufficient risk for a condition to benefit from further investigation or preventive action. Screening may be targeted to a broad group (e.g., newborn screening) or subgroups at increased risk due to some combination of age, sex, family or personal history, race/ethnicity, disease state, or other factors.

- Laboratory tests used for **diagnosis** are performed to determine the presence or absence of a specific disease or condition in symptomatic individuals. Diagnostic tests also may be used for prognosis, enabling selection of clinical care alternatives and treatments, and for monitoring treatment effectiveness and guiding treatment modifications.[19]

The regulatory policies applied to screening and diagnostic tests are similar; either may take one of two pathways to gain access to the market in the U.S.: clearance and approval by the Food and Drug Administration (FDA) or introduction under the Clinical Laboratory Improvement Amendments (CLIA). The pathway chosen typically depends on whether a test is to be marketed by a manufacturer as a product (such as in the form of a test kit) or as a laboratory developed test for use solely by the developing laboratory.

There are some key differences in the way that other policies and requirements related to clinical practice, clinical guideline development, and reimbursement (coverage and payment decisions), are applied to screening and diagnostic tests. These differences are discussed in subsequent sections of this report.

Relationship of Screening and Diagnostic Tests to Prevention

Preventive medicine has had an integral role in the growing movement toward explicit, evidence-based practice guidelines. Even so, the health system remains largely oriented to treating disease than preventing it. As such, the health and economic benefits that could accrue from wellness, prevention, and screening are far from being realized.

Three main types of prevention are primary, secondary, and tertiary. Screening and diagnostic tests contribute significantly to all three.[20] Primary prevention is aimed at preventing the onset of disease and typically involves managing risk factors in healthy people that may lead to disease (e.g., high-sensitivity C-reactive protein to detect individuals at high risk for cardiovascular events). *Secondary prevention* is aimed at treating disease after its onset but before serious complications occur (e.g., Pap test to detect and enable treatment of cervical cancer in its earliest stages). *Tertiary prevention* is used in the later or final states of a disease with the aim of minimizing the degree of disability caused by the disease (e.g., blood tests to estimate the glomerular filtration rate to monitor the severity of chronic kidney disease and manage its complications).

Many serious medical conditions can be prevented or detected early with screening and effective treatment. Evidence-based screening and preventive services have contributed to sizable reductions in morbidity and mortality in heart disease, stroke, cancer, and other major

sources of national disease burden in recent years.[21, 22] For example, routine screenings for cervical and colorectal cancer contributed to 14% decline in overall mortality from cancer (the second leading cause of death overall) during this same period.[23]

The growing body of evidence demonstrating the ability to prevent disease by screening for risk factors and adopting early intervention strategies is leading to increased use of testing for primary prevention in asymptomatic individuals. Advances in genetic and molecular testing may increasingly support the role of such tests in primary prevention through identification of individuals with predispositions for disease.[24] To date, however, clinical guidelines and payer coverage policies have focused largely on laboratory testing in secondary and tertiary prevention.[25, 26]

II. Validity and Utility of Screening and Diagnostic Tests

The development, adoption, and diffusion of laboratory tests are influenced by a widening group of stakeholders that seek well-founded evidence to support decisions about whether or how to develop technology, allow it on the market, acquire it, use it, pay for its uses, and more.[27] These stakeholders include regulators, clinicians, patients, laboratory directors, hospital managers, payers, and government policymakers, as well as manufacturers and laboratories.

Value may be assessed through clinical trials and other studies, as well as health technology assessments and other systematic appraisals of evidence that examine safety, efficacy, feasibility, effectiveness, appropriateness, and cost of tests, as well as other clinical, social, economic, and ethical implications related to their use.[28, 29] Such assessments contribute to the knowledge base for improving the quality of health care, including the development and updating of various standards and guidelines.

Assessing the value of tests may involve multiple determinations grouped into three main concepts: analytic validity, clinical validity, and clinical utility, as shown in Figure 1. While these concepts are useful in all laboratory testing, they are gaining increased attention and further definition in the realm of genetic and genomic testing. In addition to these components, cost-related assessments are of growing interest among clinicians, health care institutions, public and private sector payers, and policymakers seeking to improve value per health expenditure.[30, 31]

Figure 1. Chain of inquiry for valuation of laboratory tests

Analytic Validity	Clinical Validity	Clinical Utility
Definition: A laboratory test's ability to measure the analyte (or genotype, in the case of genetic testing) of interest accurately and reliably (i.e., the quality of the measurement)	**Definition:** A laboratory test's ability to detect and predict the disorder that is associated with an analyte measurement; a test's value to clinical decision making.	**Definition:** Clinical effectiveness; the balance of risks and benefits associated with use of a test in routine clinical practice; usefulness and value of information, positive or negative, to person being tested
Key measurements: • Accuracy • Analytic sensitivity • Analytic specificity • Precision • Robustness	**Key measurements:** • Clinical sensitivity • Clinical specificity • Positive predictive value • Negative predictive value	**Key measurements:** • Intermediate/surrogate outcomes • Health outcomes (mortality, morbidity, quality of life) • Adverse effects of diagnostic use • Adverse effects of treatment

Economic outcomes

Types of economic outcomes measurements:
- Estimates of the economic value relative to investment and may include analyses such as cost per test, patient, treatment and episode of care
- Estimates of the budget impact of a test on a provider, provider organization, health system
- Estimates the tradeoff value between costs and benefits to health are, including cost-effectiveness, cost-utility, and cost-benefit analysis

A. Analytic Validity

The overarching value of laboratory testing is the generation of objective, scientific data about patient health to inform clinical decision-making. High levels of technical performance are required to ensure quality of test results. Regarding analytic validity, laboratory test systems are designed to ensure test-to-test accuracy, precision, and robustness. Measurements aim to determine whether the test performs reliably to specifications and delivers accurate information in a laboratory setting. **Accuracy,** which is the degree of closeness of a measured or calculated quantity to its actual (true) value as compared to a reference "gold standard," includes analytic sensitivity and analytic specificity. **Analytic sensitivity** is the probability that a test will detect a specific analyte (e.g., a biomarker or genotype) when it is truly present in a specimen. **Analytic specificity** is the probability that a test will be negative when a specific analyte is truly absent in a specimen. **Precision** (or reliability, reproducibility or repeatability) is the degree to which further measurements show the same or similar results. **Robustness** refers to the resistance of test results to small changes in preanalytic or analytic variables associated with testing.[32,93]

B. Clinical Validity

Clinical validity refers to the ability of a test to detect the condition (disease or disorder) that is associated with an analyte measurement and predict the probability of having the condition based on the test result. It includes clinical sensitivity, clinical specificity (incorporating analytical validity), and positive and negative predictive values. In genetic testing, clinical validity may also be affected by factors that confound the association between a genotype and a phenotype, such as reduced penetrance (i.e., the proportion of individuals with a disease-related genotype or mutation who develop disease), variable expressivity of the disease among individuals with the same genotype, and other genetic or environmental factors.[93]

Clinical sensitivity refers to the proportion of individuals with a specified condition whose test results indicate that the condition is present, e.g., how often the genetic mutation (genotype) that is associated with the condition (phenotype) is identified in people who truly have the condition). Tests with high clinical sensitivity are useful for "ruling out" a condition if an individual tests negative. **Clinical specificity** refers to the proportion of individuals who do not have a specified condition whose test results indicate that the condition is not present. Tests with high specificity are useful for "ruling in" a condition if a person tests positive.

Positive predictive value (PPV) refers to the probability that an individual with a positive test has, or will develop, the specified condition that the test is designed to detect. **Negative predictive value (NPV)** refers to the probability that an individual a negative test result actually does not have the specified condition. Cutoff values for a test (defining presence or absence of a condition by level of the test result) represent a compromise between clinical sensitivity and clinical specificity. PPV and NPV depend not only on sensitivity and specificity but on the prevalence of a condition within the population being tested and changes in disease patterns over time. As such, PPV and NPV are not constant performance characteristics of a test. For example, if a disease is very rare in the population, even tests with high sensitivity and high specificity can have low PPV, generating more false-positive than false-negative results.[27] In these instances, without educational efforts, some providers may not fully understand the limits of a test's predictive value and may overestimate the probability of disease in patients with a positive result.[33, 34] The example of rapid testing for the influenza virus (Box 3) demonstrates the fluctuations in PPV and NPV according to prevalence in the population.

Box 3
Example: Rapid Testing for Influenza

An example of the inter-relationships between sensitivity, specificity, and predictive values can be demonstrated with **rapid testing for influenza** virus as disease patterns fluctuate from low- to mid-prevalence periods to peak flu season, then tapering back down to mid- and low-prevalence periods. Influenza viral pathogens cause significant mortality and morbidity; the recent re-emergence of a novel human influenza A virus (H1N1) posed a serious personal and public health threat.[35] Several types of rapid molecular tests have helped to detect the virus, including point-of-care tests used by physicians in clinical practice and conventional and real-time polymerase chain reaction (PCR) assays that usually have high specificity.[36] One study estimated PPV and NPV using six-year averages of weekly influenza activity reported via CDC surveillance data and a rapid molecular test with sensitivity of 70% and specificity of 90%. During a typical flu season, PPV fluctuated from 17% during low prevalence to 71% at peak while the NPV decreased from 99% during non-flu periods to 89% during flu season.[37]

The following two examples of screening for fragile X syndrome and *KRAS* gene mutation support the clinical validity of a genetic test (Box 4) and a pharmacogenomic test (Box 5) for clinical decision-making.

Box 4
Example: Fragile X Genetic Testing

Fragile X syndrome is one of the most common inherited causes of mental impairment ranging from learning disabilities to autism and severe mental retardation.[38] Because of the lack of definitive clinical diagnostic criteria, molecular tests are important for detection of individuals with fragile X.[39] The standard screening/diagnostic test is Southern blot analysis although there is increasing interest and evidence to support use of PCR-based methods. Diagnostic tests can 'rule in' or 'rule out' disease in children with developmental, speech, language, or motor delay. Individuals with a family history of fragile X can be tested to determine if they may be asymptomatic carriers of the disorder. Individuals known to be carriers also can use screening tests to determine a prenatal diagnosis of the fragile X mutation. While there is no cure for fragile X syndrome, early diagnosis facilitates clinical decision-making and therapeutic planning of interventions for speech and language, behavior, cognitive and gross motor development, sensory integration, and daily living activities, in order to improve quality of life for those with the disease.

Box 5
Example: KRAS Mutation Testing for Patients with Metastatic Colorectal Cancer

Colorectal cancer is the third most commonly diagnosed cancer and third-highest cause of cancer death for men and women in the U.S.[40] Up to 20% of patients with colorectal cancer will present with metastases, with a 5-year survival of less than 10%. Cetuximab (Erbitux®) and panitumumab (Vectibix®) are monoclonal antibodies that bind to the epidermal growth factor receptor (EGFR), inhibiting growth of metastatic colorectal cancer. However, these drugs have considerable adverse effects. Moreover, a proportion of patients with colorectal cancer have tumors with a somatic *KRAS* mutation that affects tumor response to EGFR inhibitors. Diagnostic testing for *KRAS* gene mutational status is an important predictor of non-response to EGFR-targeted therapy, with great value to clinical decision-making.[41] Retrospective analyses of data from several randomized controlled trials involving patients receiving combination cetuximab and chemotherapy demonstrated that individuals with normal (or wild-type) *KRAS* had significant improvements in tumor response and that few or none of those with mutated–type *KRAS* responded to cetuximab.[42, 43] In July 2009, the FDA announced revisions to the prescribing information for EGFR inhibitors and colorectal cancer, requiring inclusion of information on variations in the *KRAS* gene that may affect patient response to the drugs.

As noted above, clinical validity can vary with the prevalence of a condition within the population being tested, changes in disease patterns over time, and other factors, e.g., age, race/ethnicity, family history, severity of disease, and comorbidities. For many conditions, patients with severe disease are more likely to have positive tests than those with mild or early-stage disease, and healthy patients are more likely to have negative tests than those with significant comorbidities.[44, 45] For example, C-reactive protein (CRP), a sensitive marker of the acute-phase response, has been associated with future cardiovascular endpoints independently of other risk factors. A recent study of myocardial infarction survivors receiving regular high-sensitivity CRP (hsCRP) monitoring tests examined the effect of comorbidities and environmental factors on diagnostic value.[46] There were larger variations in hsCRP values for males, smokers, and patients with increased HbA1c levels >6.5%. As a result, one or two hsCRP measurements may not be sufficient to adequately characterize cardiac risk in different patient groups after myocardial infarction.

A complicating factor in examining the effect of individual patient characteristics on clinical validity is the lack of information available for these stratifications. Several reports in the literature state that, often, researchers do not consistently, or in some cases adequately, report on important patient characteristics and study design features associated with their investigations of test accuracy and predictive value.[47] Improvements in the reporting on patient characteristics could augment clinical research and guideline development for screening and diagnostic tests as well as facilitate tailored recommendations to specific patient subgroups.

C. Clinical Utility

Clinical utility refers to the evidence that use of a test is associated with improved clinical outcomes and its usefulness to patient and clinician decision-making. It encompasses effectiveness (utility in real clinical settings) or efficacy (utility in controlled settings such as clinical trials) and the net balance of risks and benefits associated with using a test in clinical practice. A test with clinical utility yields results, whether positive or negative, that provide information of value to the patient, clinician, or others involved in making decisions about management of a patient with a given condition.[48,93] In some instances in which there are no interventions or treatments available to prevent or treat disease, a test result can have clinical utility for life-planning purposes.

The specific outcomes for which tests can provide clinical utility vary by condition, but include the categories of mortality, morbidity, adverse events, and quality of life. Some biomarkers are used as intermediate or surrogate outcomes, in that they are known to be predictive of long-term health outcomes. Patient functional status, patient satisfaction, and other patient-reported outcomes, sometimes known as humanistic outcomes, are of increasing importance in patient-centered care.

One of the challenges of demonstrating clinical utility of testing is generating direct evidence of the impact of a test on health outcomes, which can require long follow-up times and is subject to the various intervening decisions and other environmental factors that can influence ultimate patient outcomes. Still, there are notable examples of such direct evidence, including that of FOBT screening for colorectal cancer, described in Box 6.

Box 6
Example: FOBT Screening for Colorectal Cancer

A 2008 systematic review conducted for the USPSTF of long-term follow-up of four RCTs of guaiac fecal occult blood tests (FOBTs) and endoscopic follow-up of positive test results found that colorectal cancer mortality was reduced 13-21% after 8-13 years of screening in two trials and that another two trials did not show mortality benefit until after 15-18 years of screening.[49] A 2007 systematic review conducted by the Cochrane Collaboration that identified reports of four RCTs and two other controlled trials involving more than 320,000 participants with follow-up ranging from 8 to 18 years concluded, based on the RCT evidence, that participants allocated to screening had a 16% reduction in the relative risk of colorectal cancer mortality.[50] Among those RCTs was one involving more than 46,000 patients, which found that cumulative mortality over an 18-year period was reduced 33% with annual testing and 21% in biennial testing.[51] In its 2008 recommendation, the USPSTF stated that "There is convincing evidence that screening with any of the 3 recommended tests [fecal occult blood testing, sigmoidoscopy, or colonoscopy] reduces colorectal cancer mortality in adults age 50 to 75 years." The USPSTF noted the importance of considering the lag time between testing and being able to demonstrate changes in outcomes: "There is adequate evidence that the benefits of detection and early intervention decline after age 75 years. The lead time between the detection and treatment of colorectal neoplasia and a mortality benefit is substantial, and competing causes of mortality make it progressively less likely that this benefit will be realized with advancing age.[52]

In many instances, the impact of laboratory testing on health outcomes is inferred through linkages between the tests and intermediate outcomes (or endpoints) or surrogate outcomes that are known to be strongly associated with long-term health outcomes.[e] Surrogate outcomes, such as HbA1c for diabetes morbidity and CD4 cell counts and viral RNA levels for progression of AIDS, are generally easier to measure than the long-term health outcomes with which they are associated and, thus, are used more commonly to determine the effects of testing and other health care interventions.[8] These associations must be validated in appropriately long-term natural history studies or other appropriate study designs. Some surrogate outcomes are better predictors of patient outcomes than others. Examples of other validated ones include: cardiac troponin for identifying and predicting outcome of acute coronary syndrome, prostate-specific antigen (PSA) levels for predicting the incidence and course of prostate cancer, INR/prothrombin time for predicting thrombolytic events, and HER-2/*neu* for predicting response to treatment with trastuzumab (Herceptin) in women with breast cancer.

1. *Comparative Effectiveness Research*

Recent efforts to augment the national capacity for conducting and using comparative effectiveness research (CER) will increase the interest in such research, particularly for clinical utility, of clinical laboratory testing. As is clear in recent reports by the IOM and the Federal Coordinating Council for Comparative Effectiveness Research on recommended national priorities for CER, this research will address the full range of interventions, including drugs,

---

[e] A biological marker (biomarker) is an objectively measured characteristic that is used as an indicator of normal biological processes, natural history of disease, or effect of a therapy. A clinical endpoint is a characteristic or variable that reflects how a patient feels, functions, or survives. An intermediate (nonultimate) endpoint is a true clinical endpoint (e.g., a symptom or measure of function, such as symptoms of angina frequency or exercise tolerance) but not the ultimate outcome of the disease (e.g., survival or the rate of other serious and irreversible morbid events). A surrogate endpoint is intended to substitute for or predict a clinical endpoint on the basis of epidemiological, therapeutic, pathophysiologic, or other scientific evidence.[53]

biologics, tests, imaging, and medical and surgical procedures, as well as health care delivery system organization, delivery, and financing.

While there is no standard definition of CER, most address a combination of the following attributes or emphases:

- Direct comparisons of alternative interventions (as opposed to comparison with placebo or indirect comparisons)

- Effectiveness (in realistic health care settings) rather than efficacy (in ideal circumstances)

- Health care outcomes (e.g., morbidity, mortality, QoL, adverse events, and symptoms) rather than intermediate or surrogates endpoints

- Use of primary and secondary data collection, with emphasis on head-to-head comparisons in RCTs and practical/pragmatic clinical trials as well as observational studies (using registries, claims data, electronic health records) and systematic reviews

The definition used by the IOM is:

*[T]he generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels. CER's distinguishing characteristics include informing a specific clinical or policy decision, comparing at least two approaches or interventions, describing results at the subgroup level, measuring benefits in real-world populations, and applying appropriate methods and data sources. —*Institute of Medicine[54]

Two main ways in which laboratory testing will be involved in CER are as an indicator of intermediate and long-term health outcomes in comparisons of other interventions and as the index (focal) intervention in a direct comparison with alternatives (Box 7). Pursuant to involvement as index interventions, among the top tier of 25 CER priorities recommended by the IOM in its June 2009 report were:

- Compare the effectiveness of various screening, prophylaxis, and treatment interventions in eradicating methicillin resistant *Staphylococcus aureus* (MRSA) in communities, institutions, and hospitals.

- Compare the effectiveness of genetic and biomarker testing and usual care in preventing and treating breast, colorectal, prostate, lung, and ovarian cancer, and possibly other clinical conditions for which promising biomarkers exist.

Box 7
Main Roles of Clinical Laboratory Tests in CER

- **As an indicator of patient intermediate and long-term outcomes in comparisons of other treatments/interventions,** particularly those related to priority health conditions, such as diabetes, obesity, heart disease, stroke, kidney disease, HIV/AIDS, mental health/substance abuse, pneumonia, cervical and colon cancer, and pregnancy. For example, HbA1c could be used as one indicator to evaluate the effectiveness of alternative oral medication regimens (e.g., sulfonylureas, biguanides, dipeptidyl peptidase IV inhibitors, or combination therapy) to treat type 2 diabetes. Laboratory values can be useful in quantifying baseline characteristics, assessing intermediate outcomes, conducting subgroup analysis, and more.

- **As the subject of analysis for head-to-head comparisons of alternative laboratory tests or comparisons of a laboratory test to another test intervention for a particular health care condition.** For example, a CER study may compare a new molecular test assessing the gene expression profile of RNA from the peripheral blood of heart transplant recipients (i.e., AlloMap HTx based on blood draws) to the current standard of care assessing RNA from invasive endomyocardial (heart tissue) biopsy procedures.[55] Another CER study might evaluate the effectiveness of HPV testing for cervical cancer compared to conventional Pap smear testing.[56]

---

2. *Challenges of Demonstrating Causal Effect of Testing on Outcomes*

The size of the literature documenting clinical utility of laboratory testing, including patient outcomes, is small relative to the corresponding literature on analytical validity and clinical validity. Typically, there are substantial difficulties in establishing direct causal links between ordering a test and changes in mortality, morbidity, quality of life, and other major patient health outcomes, as a test is likely to be just one of many interventions and environmental and behavioral determinants of patient outcomes.[57] In order to influence outcomes, a laboratory test must be ordered, conducted, returned with results on a timely basis, appropriately interpreted, and affect a decision for further diagnosis or treatment that results in changes in outcomes. This includes instances in which laboratory testing is used to monitor intermediate outcomes of treatment in order to guide the next treatment decision toward desirable outcomes.

As part of the larger clinical process, laboratory testing can affect other aspects of patient care, such as timing, efficiency, and patient and provider satisfaction.[58, 59] Providers may interpret and act on laboratory test results differently and unpredictably, which can affect patient outcomes.[60] Providers and patients may ignore positive results or proceed with a clinical intervention despite a negative test.

Several factors can confound the ability to conduct population-based laboratory outcomes studies. These include the lack of standardization of data collection and reporting methods, lack of agreement regarding appropriate analysis of data (e.g., whether or not to risk-adjust data), and the high cost of collecting outcomes data. Outcomes measurement can be severely constrained by sample size, missing or incomplete test results in patients' medical records, limited ability to perform risk adjustment with data abstracted from administrative records, and higher cost to abstract data from medical records.[61] Other challenges can include the inability to conceal the identity of tested versus non-tested patients (in blinded RCTs), the number of patients or volunteers required for a study to achieve statistical significance, and what can be long periods of time between a test and a patient outcome of interest, as suggested by the example of screening for colorectal cancer noted above.[58, 60, 62, 63] For these reasons, the assessment of the impact of

laboratory tests on health outcomes has relied more often on intermediate or surrogate outcomes, computer simulations (modeling), and use of observational studies and other non-experimental study designs. However, several trends should improve the capacity for conducting outcomes studies of clinical laboratory testing. These include the current emphasis on evidence-based decision-making, advances in data mining and other techniques for analyzing claims data and other administrative and observational data sets, expansion of electronic health records systems and networks, and the increase in funding for CER and related methods and infrastructure development.

D. Economic Outcomes and Impact

Interest in cost-effectiveness analyses and related studies of the economic impact of health care technologies continues to increase among stakeholders. In screening and diagnostic testing, the body of such evidence is small, though growing.[30, 31] Recent debate about whether savings can be realized from greater federal investment in preventive and wellness services and their cost-effectiveness relative to other types of health care highlights the importance of rigorous, policy-relevant demonstrations of the economic impact of laboratory testing used in screening and diagnosis.[64] The main types of analysis and their application to testing are summarized below. While these analyses are used more widely and various expert groups and journals are working to improve their quality, considerable variation persists in the methods, results, and reporting of these analyses.

Cost-of-illness analysis is a determination of the economic impact of an illness or condition (typically on a given population, region, or country) e.g., of heart disease, diabetes, or cancer, including associated treatment costs. Cost-of-illness studies are used to estimate the burden of disease and provide an economic context for opportunities to lowering such burdens with health care interventions. For example, Box 8 describes how cost-of-illness provides the context for potential savings from MRSA testing.

Box 8
Example: Cost-of-Illness Analysis of Rapid MRSA Testing

The economic burden of health care-acquired infections is substantial. Adjusted to 2007 dollars, CDC estimates that the direct cost per case for all HAIs ranges from $20,549 to $25,903, for a total of roughly $36 to 45 billion in annual costs.[65] Of these costs, $3.5 to 10 billion are associated with surgical site infections; $670 million to $2.7 billion to central venous line associated bloodstream infection, $1 to $1.5 billion in ventilator-associated pneumonia, and $1.2 to $1.6 billion catheter-associated urinary tract infections. Approximately 50% of all HAIs are MRSA-related.[66, 67]

A nine-year study at Brigham and Women's Hospital (Boston) found that routine surveillance cultures and subsequent contact precautions decreased incidence of bacteremia by 75% in ICUs, 40% in non-ICU areas, and 67% hospital-wide.[68] After issuing a mandate for MRSA testing in all high-risk units, the VA reduced infection rates in ICUs by 79% at their Palo Alto site.[69, 69] This large reduction in rates of MRSA infection rates at a major facility suggests the potential for broader-scale dramatic reductions in the prevalence of MRSA and the morbidity, mortality, and costs associated with that pathogen.

Cost-minimization analysis is a determination of the least costly among alternative interventions that are assumed to produce equivalent outcomes. Several studies have examined the cost of POCT relative to central laboratory testing using metrics such as cost-per-test (including estimates for the

cost of equipment, supplies, labor, and other variables) and laboratory test turnaround time.[70, 71] Box 9 provides an example of cost-minimization analysis for bedside glucose testing.

Box 9
Example: Cost-Minimization Analysis of Bedside Glucose Testing

A study reported in 2004 compared the analytical costs of central laboratory glucose testing and semi-automated bedside glucose testing (BGT) among 445 hospitals enrolled in the College of American Pathologists' Q-Probes quality assurance program. Results showed different distributions of costs across three main types of sites. The median (10th-90th percentile range) analytical costs per glucose test were $1.18 dollars ($0.36-$5.59) for central laboratories, $1.96 ($0.77-$9.51) for high-volume BGT sites, and $4.66 ($1.02-$27.54) for low-volume BGT sites.[72] In addition to being higher than costs for central laboratories, costs for BGT were highly variable and dependent on volume. In this instance, the cost-per-test for central laboratories was the lowest, although ranges of medians showed considerable overlap. Of further consideration is the potential clinical value to the provider and patient in having an instantaneous (bedside) result, which could lead to cost savings if it avoids helps to avoid unnecessary care.

Cost-effectiveness analysis (CEA) compares costs in monetary units with outcomes in quantitative non-monetary units, e.g., reduced mortality or morbidity. CEA aims to weigh the health and economic tradeoffs of alternative health interventions. It calculates the incremental (marginal) cost per incremental unit of effectiveness achieved through use of an intervention versus the standard of care or other alternative. Results are presented as net cost per health outcome, such as cost per case prevented or cost per life saved; this is also known as an incremental cost-effectiveness ratio (ICER).[73, 74] One type of cost-effectiveness analysis, sometimes referred to as **cost-utility analysis (CUA)**, compares costs in monetary units with outcomes in terms of their utility, usually to the patient, measured, e.g., in quality-adjusted life years (QALYs), a unit combining quality of life and length of life.[74] Although U.S. payers use no formal threshold for an acceptable cost per QALY, incremental cost-effectiveness ratios of $50,000-$100,000 or less per QALY are generally regarded as acceptable value.[75] CEA/CUA can be conducted from different economic perspectives, e.g., of the clinician, payer, patient, or society at large. Various types of cost-effectiveness analyses have been conducted in such areas as prenatal genetic screening, screening for *Chlamydia trachomatis*, FOBT for colon cancer, HER-2 /*neu* testing of breast cancer, HIV screening, and nucleic acid testing for safety of donated blood.[76-81] An example is highlighted in Box 10.

Cost-benefit analysis (CBA) compares costs and benefits, both of which are quantified in common monetary units. Two basic approaches for cost-benefit analysis (CBA) are ratio approach and the net benefit approach. The ratio approach indicates the amount of benefits (or outcomes) that can be realized per unit expenditure on a technology vs. a comparator. In the ratio approach, a technology is cost beneficial vs. a comparator if the ratio of the change in costs to the change in benefits is less than one. The net benefits approach indicates the absolute amount of money saved or lost due to a use of a technology vs. a comparator. In the net benefits formulation, a technology is cost-beneficial vs. a comparator if the net change in benefits exceeds the net change in costs. Due largely to the difficulty associated with assigning monetary values to years of life or health outcomes, few true CBAs have been conducted for health care technologies, including for laboratory testing.

Budget impact analysis (BIA) provides an estimate of the financial impact on capital and operating budgets (and may include analysis of cost offsets) of the adoption and use of a health care intervention in a given patient population, health care setting, or other context of the introduction of a technology or service. BIA often includes analysis of cost offsets. CEA and BIA can provide complementary analyses for evaluating technology acquisition and use decisions by health care providers, policy makers, and others.

Box 10
Example: Cost-effectiveness Analysis of Factor V Leiden Test for Thrombophilia

Thrombophilia is the propensity to develop potentially venous thromboembolism (VTE) caused by hereditary defects in one or more of the clotting factors. The most common mutation is factor V Leiden associated with activated protein C (APC) resistance (an anticoagulation enzyme). The factor V Leiden mutation has a relatively high prevalence in the general population, including 5% in Caucasians, and accounts for 85-95% of APC resistance cases. Patients at high-risk for VTEs include those using oral estrogen preparations, who are pregnant, or having major surgery. Increased screening for the factor V Leiden mutation has been suggested as a strategy for preventing VTEs. The factor V Leiden test is a genetic test for detecting the presence of the mutation.

A recent systematic review and CEA conducted in the U.K. examined the cost effectiveness of screening regimens for four types of high-risk individuals.[82] The cost-effectiveness ratio was expressed as costs per adverse clinical complication prevented when comparing universal screening and selected screening to no screening for each of the four risk groups. The study was based on a hypothetical model of 10,000 patients in each of group. The results for universal screening versus selective screening for each of the four groups were as follows. Shown are the number of complications prevented by each strategy and the ICER, i.e., the ratio of the incremental cost for each strategy to the number of complications prevented.[82]

	Universal Screening		Selective Screening	
	Complications Prevented	*ICER**	*Complications Prevented*	*ICER*
Combined oral estrogen:	3	£200,402 ($323,230)	1	£79,085 ($127,560)
Hormone replacement therapy	42	6,824 (11,010)	15	2,446 (3,950)
Pregnancy	59	81,554 (131,540)	7	81,250 (131,030)
Orthopedic surgery	88	14,129 (22,790)	26	9,136 (14,740)

*ICER: Cost in UK£ (and US$) per complication prevented

The analysis offers important tradeoffs for consideration. First, for all four risk groups, universal screening led to prevention of more complications than did selective screening, e.g., 42 vs. 15 in the hormone replacement therapy risk group. Second, for all four risk groups, selective screening yielded lower ICERs than universal screening, demonstrating better cost-effectiveness for that strategy in general. The greatest improvements in cost-effectiveness resulting from selective screening were in the combined oral estrogen risk group (61% decrease) and the hormone replacement therapy risk group (64% decrease). Third, despite their relative magnitudes, the ICERs of both universal and selective screening of the hormone replacement therapy group were the lowest among the full set of eight strategies.

In its August 2009 letter to Congress about the potential economic impact of greater federal spending on preventive and wellness services, the CBO focused on how it would "score" such spending. Scoring by the CBO for determining the net impact of proposals to expand governmental support for a designated program is a form of BIA. CBO stated that, although

particular types of preventive care would have different effects on costs, available evidence suggests that expanded use of most preventive services results in higher, not lower, overall spending.[f] CBO also distinguished its scoring from determinations of the cost-effectiveness of spending on such services, noting that, while most would add to spending, many would be cost-effective, i.e., their benefits to health would be justified by the additional net costs.[64]

Further deliberations on these matters in the context of national health reform and other proposals to improve the U.S. health care system should address how to assess the economic impact of these services. This should include, for example, the extent to which CBO scorekeeping rules and procedure are limited with respect to the broader costs and benefits (i.e., beyond federal budgetary effects) that would accrue to the nation, how some preventive services (or selective use of them in certain high-risk populations) are more likely to be cost saving or cost effective than others, and various methodological aspects such as the time horizon of analysis and inclusion of direct and indirect costs.

As is so for many other types of health care technology, greater investment is needed to build the body of evidence on the economic impact of laboratory testing. Findings about laboratory tests that are truly cost-saving or highly cost-effective, such as colorectal cancer screening in adults aged 50-75 and screening young women for chlamydial infection,[22, 83] should be broadly recognized and used as models for further work. Among key considerations in such analyses are that the economic impacts of specific tests can vary significantly depending on the populations targeted (with testing of high-risk populations more likely to be cost-effective) and tradeoffs between the cost of greater testing frequency and yield of cases detected.

III. Analytic Frameworks to Appraise Evidence of Clinical Value

Health care decision-makers increasingly rely on health technology assessments (HTAs), systematic reviews, and other evidence reports to support decisions and policies related to clinical care, practice guidelines, coverage policies, public health services, and other purposes.[27] Usually, these evidence reports are prepared after clinical laboratory tests have received regulatory approval by the FDA or met specifications of CLIA, as appropriate. Stakeholder organizations may undertake or sponsor their own HTAs for various reasons, such as to support decisions regarding health plan benefits, clinical practice guidelines, technology acquisition, public health services, coverage and payment, and investment in technology.[27]

HTAs pertaining to testing are often organized according to analytic frameworks such as the one shown in Figure 2. These analytic frameworks map the linkages between the populations to be tested, test results, interventions (performance of a test, treatment), change in intermediate outcomes, and change in patient outcomes.[84] A set of specific key questions are developed to provide the conceptual context for the study and guide retrieval of the available scientific evidence to assess these linkages.

---

[f] Among the points made by CBO with regard to cost offsets is that some of the proposed additional federal spending on these services would only substitute for existing spending on these services from other sources, resulting only in a cost shift to the federal government with no additional benefits.

A. Organizations that Conduct Evidence Appraisals

Many different organizations conduct HTAs or other evidence reports of health technologies. Some organizations evaluate technologies for the spectrum of clinical applications, including screening, diagnosis, treatment, and monitoring, whereas others limit their evaluations to certain types of interventions or services.

The **U.S. Preventive Services Task Force (USPSTF)** is an independent, non-federal advisory group sponsored by the Agency for Healthcare Research and Quality (AHRQ) that conducts evidence reviews of screening and preventive services used in primary care.[84, 85] The panel of experts that serve on the USPSTF evaluates the strength of the scientific evidence and makes recommendations for use of preventive services based on its findings.[25] The type of clinical analytic framework used by the USPSTF is recognized as a standard approach for mapping linkages between patient populations, interventions and outcomes, shown in Figure 2.

The Evaluation of Genomic Application in Practice and Prevention (EGAPP) initiative, sponsored by CDC, conducts rigorous evaluations of genetic tests and other genomic applications for clinical and public health practice in the U.S.[86] An independent, non-federal advisory group, EGAPP currently focuses on evaluation of tests with wide population application (e.g., high disorder prevalence, higher frequency of test use), those with potential to affect clinical and public health practice (e.g., emerging prognostic and pharmacogenomic tests), and those for which there is significant demand for information. EGAPP has an evidence grading system that is based on hierarchies for each of analytical validity, clinical validity, and clinical utility (described below).

The Community Services Task Force, sponsored by the Centers for Disease Control and Prevention (CDC), conducts evidence reviews to develop guidelines for screening and prevention for services related to infectious diseases and other public health concerns, e.g., health care-acquired infections such as *S. aureus* (staph infection), sexually transmitted diseases such as *Chlamydia trachomatis* and HIV/AIDS, and potential epidemic and pandemic diseases such as influenza.

Through its **Effective Health Care Program**, including its 14 Evidence-based Practice Centers (EPCs), located primary at academic health centers, and other programs, **AHRQ** funds systematic evidence reviews of relevant scientific literature on clinical, behavioral, and organizational topics. These are used by a wide array of stakeholders to inform the development of educational materials and tools, quality measures, guidelines, coverage decisions, and research agendas. In particular, the Oregon EPC, based at Oregon Health & Science University, supports the work of the USPSTF, and the EPCs also support EGAPP. Although these particular EPC evidence reviews are prepared initially for USPSTF and EGAPP, they are placed in the public domain and are used by others. In addition to biomedical research that supports the development and validation of many and diverse tests, the **National Institutes of Health,** including the National Heart, Lung, and Blood Institute and the National Cancer Institute, undertake studies to support development of clinical guidelines on use of screening and diagnostic tests.

Health professional societies, particularly those associated with medical specialty areas, have guideline development committees and other expert advisory groups and panels that conduct independent evaluations of scientific evidence and make recommendations. Examples of the many such organizations engaged in these activities are the American Academy of Family Physicians, American Academy of Pediatrics, American College of Cardiology, American

College of Physicians, and American Society of Clinical Oncology. Non-profit associations devoted to certain health problems, such as the American Cancer Society, American Diabetes Association, and American Heart Association, are also active in sponsoring advisory groups and expert panels that evaluate evidence and develop guidelines related to testing. Nearly 300 health professional organizations have clinical guidelines included in the National Guideline Clearinghouse, managed by AHRQ.[g]

Many **private sector payers** have internal research arms that undertake assessments to determine coverage and reimbursement. Examples of major commercial health plans that conduct formal reviews of new technologies include Aetna, CIGNA, UnitedHealthcare, and WellPoint.[14] HTA vendors such as the Cochrane Collaboration, BlueCross BlueShield Association Technology Evaluation Center (TEC), ECRI Institute, and Hayes, Inc.,[h] produce HTAs of tests using in screening and diagnosis.[87, 88] BlueCross BlueShield TEC and ECRI Institute are also EPCs.

Figure 2. Generic Clinical Analytic Framework for Screening and Diagnostic Tests
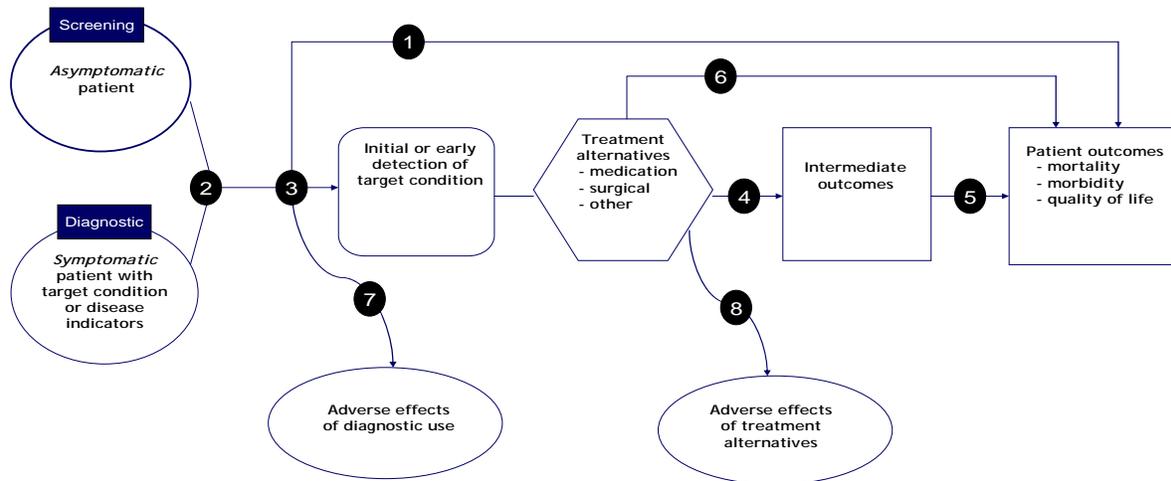
Key questions correspond to numbers above.

(1) Is there direct evidence that the test reduces morbidity, morbidity, and/or quality of life?

(2) What is the prevalence of disease in the target group? Can a high-risk group be reliably identified?

(3) Can the test accurately detect the target condition? (a) What are the sensitivity and specificity of the test? (b) Is there significant variation between examiners in how the test is performed? (c) In actual testing programs, how much earlier are patients identified and treated?

(4) Does treatment reduce the incidence of the intermediate outcome? (a) Does treatment work under ideal, clinical trial conditions? (b) How do the efficacy and effectiveness of treatments compare in community settings?

(5) Is the intermediate outcome reliably associated with reduced morbidity and/or mortality?

(6) Does treatment improve health outcomes for people diagnosed clinically? (a) How similar are people diagnosed clinically to those diagnosed by screening? (b) Are there reasons to expect people diagnosed by screening to have even better health outcomes than those diagnosed clinically?

(7) Does testing result in adverse effects? (a) Is the test acceptable to patients? (b) What are the potential harms, and how often do they occur?

(8) Does treatment result in adverse effects?

---

[g] A complete listing of the organizations can be found on the National Guideline Clearinghouse website at http://www.guidelines.gov/browse/browseorgsbyLtr.aspx?Letter=*.

[h] Hayes, Inc., provides technology assessment reports for health plans, managed care companies, hospitals, and health networks and offers training programs to facilitate participants' understanding of the HTA process.

B. Direct and Indirect Evidence of Test Impact on Outcomes

The analytic framework of the type shown in Figure 2 and used by such groups as the USPSTF and EGAPP provides routes for evaluating direct and indirect evidence of the impact of testing on patient outcomes. The direct route (#1 in Figure 2) follows a population that has been tested long enough to determine whether the test had any ultimate impact on specified patient outcomes, e.g., of improvements in mortality, morbidity, or quality of life. Prospective experimental study designs, particularly RCTs, are the most rigorous for assessing test impact along this direct route. As discussed below, generating direct evidence in this manner can be challenging in many instances, and such studies often are not available.

When direct evidence of the impact of testing on patient outcomes using RCTs or similarly rigorous methods is unavailable, impact on outcomes may be established through a chain of indirect evidence, also shown in Figure 2. This would include assembling evidence from separate studies establishing the ability of the test to detect a target condition (#3) and a treatment choice that affects intermediate outcomes (#4) or patient outcomes directly (#5), and evidence that these intermediate outcomes are strongly associated with patient outcomes (#6). Separate studies can be used to assess adverse effects of the test (#7) or of the treatments (#8). Often at issue is whether such indirect chains of evidence are sufficient to provide expert panels enough confidence to conclude that a test has an impact on outcomes and can be recommended for clinical use.

C. Evidence Hierarchies Used for Screening and Diagnostic Testing

HTAs and other evidence reports appraise different types of studies in the available bodies of evidence pertaining to the interventions being assessed. The various study designs offer characteristic strengths and limitations for answering evidence questions.[84, 85] In most instances, whether for evidence reports sponsored by government, private non-profit, or commercial programs, determinations of the methodological quality of the available studies are based on evidence hierarchies and related quality criteria. Most of these hierarchies are arranged according to the relative rigor of various study designs to account for the internal validity of the causal effect of an intervention on specified outcomes. Typically, these hierarchies rank RCTs, or systematic reviews of RCTs or meta-analyses of RCTs, as the highest form of evidence.[89] Although systematic reviews (some of which include meta-analyses) and meta-analyses do not themselves constitute primary evidence, they are compilations or syntheses of evidence that may hold greater weight than the individual RCTs (or other primary studies) of which they are comprised.

Two evidence hierarchies of particular importance for laboratory tests used in screening and diagnosis are those of the USPSTF and EGAPP, shown in Table 2 and Table 3, respectively. The USPSTF and EGAPP account not only for study design but for the quality of study execution, e.g., "properly powered and conducted RCTs" and "well-designed longitudinal cohort studies." In appraising the strength of evidence for clinical preventive services, the USPSTF places RCTs and systematic reviews or meta-analyses of RCTs at the top level. EGAPP has evidence hierarchies for each of analytic validity, clinical validity, and clinical utility. EGAPP places meta-analysis of RCTs at the highest level, followed by single RCTs. In addition to its evidence hierarchies, the USPSTF has a grading system for its recommendations, shown in Table 4 and Table 5. EGAPP has a similar system (not included here).

Table 2. USPSTF Hierarchy of Research Design

Level	Evidence Type
I	Properly powered and conducted randomized controlled trial (RCT); well-conducted systematic review or meta-analysis of homogeneous RCTs
II-1	Well-designed controlled trial without randomization
II-2	Well-designed cohort or case-control analytic study
II-3	Multiple time series with or without the intervention; dramatic results from uncontrolled Experiments
III	Opinions of respected authorities, based on clinical experience; descriptive studies or case reports; reports of expert committees

Source: U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF. July 2008.

Despite its traditional emphasis on evidence from RCTs, the USPSTF has become somewhat more flexible in its approach to grading of evidence and recommendations, as described in its methods revisions made during 2007-08. The USPSTF's approach may benefit further by continuing to consider alternative evolving methods for evidence appraisal, including those described below. In formulating their recommendations, both USPSTF and EGAPP rely not only on strength or quality of evidence but on the anticipated magnitude of net benefit. Weighing both enables providing favorable recommendations even when the evidence is not of the highest level. For example, as is apparent in Table 4, the *certainty* of net benefit (i.e., benefit minus harms based on available evidence) of a preventive intervention may be only moderate, due, for instance, to having level II-2 or II-3 evidence rather than level I evidence. However, if the expected *magnitude* of net benefit is high or moderate, USPSTF can assign a "B" grade recommendation, meaning that the USPSTF recommends offering or providing the intervention, as indicated in Table 4.

Table 3. EGAPP Hierarchies of Data Sources and Study Designs for Components of Evaluation

Level	Analytic Validity	Clinical Validity	Clinical Utility
1 (highest)	Collaborative study using a large panel of well-characterized samples Summary data from well-designed external proficiency testing schemes or interlaboratory comparison programs	Well-designed longitudinal cohort studies Validated clinical decision rule	Meta-analysis of RCT
2	Other data from proficiency testing schemes Well-designed peer-reviewed studies (e.g., method comparisons, validation studies) Expert panel reviewed FDA summaries	Well-designed case-control studies	A single RCT
3	Less well designed peer-reviewed studies	Lower quality case-control and cross sectional studies Unvalidated clinical decision rule	Controlled trial without randomization Cohort or case-control study
4	Unpublished and/or non-peer-reviewed research, clinical laboratory, or manufacturer data Studies on performance of the same basic methodology, but used to test for a different target	Case series Unpublished and/or non-peer-reviewed research, clinical laboratory, or manufacturer data Consensus guidelines Expert opinion	Case series Unpublished and/or peer-reviewed studies Clinical laboratory or manufacturer data Consensus guidelines Expert opinion

Source: Teutsch SM, Bradley LA, Palomaki GE, et al. The Evaluation of Genomic Application in Practice and Prevention (EGAPP) initiative: methods of the EGAPP Working Group. Genet Med 2009;11(1):3-14.

Table 4. USPSTF Recommendation Grid: Letter Grade of Recommendation

Certainty of Net Benefit	Magnitude of Net Benefit			
	Substantial	Moderate	Small	Zero/Negative
High	A	B	C	D
Moderate	B	B	C	D
Low	Insufficient			

Grade A indicates that the certainty of evidence is high that the magnitude of net benefits is substantial.

Grade B indicates that the certainty of evidence is moderate that the magnitude of net benefits is either moderate or substantial, or that the certainty of evidence is high that the magnitude of net benefits is moderate.

Grade C indicates that the certainty of the evidence is either high or moderate that the magnitude of net benefits is small.

Grade D indicates that the certainty of the evidence is high or moderate that the magnitude of net benefits is either zero or negative.

Grade I indicates that the evidence is insufficient to determine the relationship between benefits and harms (ie, net benefit).

Source: U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF. July 2008.

Table 5. What the USPSTF Grades Mean and Suggestions for Practice

Grade	Grade Definitions	Suggestions for Practice
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer/provide this service.
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer/provide this service.
C	The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is moderate or high certainty that the net benefit is small.	Offer/provide this service only if there are other considerations in support of the offering/providing the service in an individual patient.
D	The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.	Discourage the use of this service.
I Statement	The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality or conflicting, and the balance of benefits and harms cannot be determined.	Read "Clinical Considerations" section of USPSTF Recommendation Statement. If offered, patients should understand the uncertainty about the balance of benefits and harms.

Source: U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF. July 2008.

In addition to periodic refinement of evidence hierarchies and grading systems used by the USPSTF and EGAPP, other systems are being developed and adopted that have components that are devoted in particular to interventions used in screening and diagnostics. Of particular note are the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system[90] and the Strength of Recommendation Taxonomy (SORT) system.[91] These systems provide explicit differentiation, including separate evidence hierarchies, of evidence requirements for screening and diagnosis, and related guidance about using indirect evidence

to make inferences about impact of these interventions on patient outcomes. GRADE is being used (sometimes in adapted forms) by an increasing number of organizations, including AHRQ EPCs, the Cochrane Collaboration, and certain medical professional societies and journals. Developed by journal editors and other experts, SORT emphasizes a patient-centered approach to appraising evidence.

D. Challenge of Relying on RCTs

The strength of RCTs is in their internal validity — how well the study design and execution diminish the opportunity to introduce bias or inferential error that can affect results regarding the true impact of interventions on outcomes under the particular conditions of the study.[92] In RCTs, randomization ensures that intervention groups (e.g., one getting a new intervention and one getting standard care) differ only in their exposure to the intervention, so that differences in observed impacts can be attributed to differences in the intervention.

Notwithstanding their strengths, there are disadvantages to RCTs in the context of evidence appraisals, particularly for laboratory tests. Given the multiple intervening steps between accurate diagnostic results and improved health outcomes, it can be time-consuming, complex, costly, and, in certain instances, not feasible, to conduct RCTs on laboratory tests that yield direct evidence on patient outcomes.[27, 93] The time between the use of a laboratory test until the time that patients experience a change in health outcomes that may have been influenced by that test can be years or even decades. Blinding of patients (as well as clinicians and investigators) in an RCT, e.g., as to whether patients received one test vs. another, or a test vs. no test, may be impractical. It may be difficult to account for all confounders that may affect patient outcomes in an RCT (e.g., "washing out" previous or current other treatments that may affect outcomes). The impact of a test may be difficult to isolate relative to the impacts of other diagnostic and therapeutic interventions, as well as any relevant environmental and behavioral factors experienced by a patient, between the time of the test and assessment of ultimate patient outcomes.[57] Results of the RCTs, typically conducted with carefully selected patients under carefully controlled conditions, may not be generalizable to the broader patient population in real-world clinical practice. RCTs may not be able to consider or account for patient preferences, which may reflect on their relevance for patient-centered care.[93, 94] To the extent that a substantial body of evidence from other study designs has accumulated, it may be strong enough to obviate the need for RCTs. Further, enrolling sufficient numbers of patients in RCTs and identifying clinical investigators may be impractical under those circumstances.

Methods other than RCTs are better suited to answering certain questions about the clinical impact or value of health care interventions, including screening and diagnostic testing. For example, the best method for assessing prognosis of a condition (disease, disorder) may be a patient cohort study of people with the condition with follow-up at uniform time intervals in the clinical course of the condition. Case control studies can be used to identify risk factors for a condition. Test accuracy (sensitivity, specificity, positive and negative predictive value) can be assessed by a large cross-sectional study of patients known to have a condition. Registries and surveillance studies can be used to monitor patient populations for the incidence of serious or rare adverse effects that may not arise in RCTs that are too small, insufficiently representative of target patient populations, or too short in duration to detect such adverse effects. Although clinical trials that are not randomized and observational studies tend to be less rigorous (for

minimizing sources of bias) than RCTs, a well-designed studies of these types can provide evidence that is sufficiently strong for making clinical and policy decisions.[23, 85, 27, 95]

As noted above, evidence requirements and grading approaches that are applied to laboratory testing are gradually changing and are being affected by trends in generation of primary data and analyses of observational data. While RCTs will remain the preferred study design for establishing causal effects of interventions on patient outcomes, strengthening of other study designs and emergence of others are helping to supplement and, in some instances, substitute for RCTs. Included are evolving variations in traditional clinical trial designs (e.g., practical clinical trials and adaptive and Bayesian trial designs) and "mining" and other analyses of clinical data sources (e.g., electronic medical records, patient registries) to evaluate populations, interventions, and outcomes. New methods also may involve studies by clinical laboratories that can link de-identified patient samples gathered for testing purposes to patient outcomes. New methods and tools for developing and evaluating evidence are necessary to adequately address the effectiveness of interventions on risks, the changing disease patterns of comorbidities, and heterogeneity of treatment effects based on individuals' genetic variations.[96]

E. Challenges Associated with Implementation of Screening in Clinical Practice

1. High Value Screening Tests Underused by Clinicians

There is substantial evidence that clinicians often fail to order appropriate clinical laboratory tests to screen for, diagnose, and monitor patient health conditions, including those tests recommended in clinical guidelines or incorporated into performance measures to evaluate quality of care.[97, 98] Underuse is a common problem for several laboratory tests that reduce mortality by detecting disease early or that prompt interventions to control risk. For example, the American Cancer Society recommends that, beginning at age 50, men and women at average risk for developing colorectal cancer receive (depending on patient preference and other factors) one of a set of tests that detect polyps and cancer (flexible sigmoidoscopy, colonoscopy, double-contrast barium enema, or CT colonography) or that mainly find cancer (fecal occult blood test [FOBT], fecal immunochemical test [FIT], or stool DNA test [sDNA]) at designated intervals, including follow-up colonoscopy when certain of these tests are positive.[99] Yet, the 2006 National Healthcare Quality Report (NHQR) reported that only 52% of adults over age 50 had colorectal cancer screening by any method in the previous 2 years, indicating that just under half of adults in this population are not following the recommended screening schedule.[100] Some tests are overused relative to existing guidelines; one example concerns Pap smears. A study published in 2004 reported that half of all women who had undergone hysterectomy (or nearly 10 million women) received unnecessary Pap smears within the three years leading up to 2002, excluding those Pap smears that may have preceded a recent hysterectomy and hysterectomies that spared the cervix or were performed for cervical neoplasia.[101]

Clinical laboratory tests also are important tools to control, manage, and monitor chronic conditions. They can detect complications of care and prevent the development of additional comorbidities. Underuse also is prevalent in this regard. For example, elevated blood cholesterol (i.e., LDL and total cholesterol) is an especially important risk factor for heart disease and contributes to the management of diabetes. Significant progress has been made in raising patients' awareness of the importance of cholesterol screening. However, according to an analysis by the National Center for Health Statistics, only 65% of men and 70% of women 20

years or older have had their cholesterol level checked in the last five years.[102] According to the NHQR, the lack of adequate screening contributes to overall poor control of high cholesterol in the population.[103] Of course, measurement alone does not assure desired outcomes; rather, patients must receive needed therapeutic adjustments and counseling from clinicians for effective metabolic control.[104] Nevertheless, without reliable test results, the clinician cannot make appropriate adjustments or give informed medical advice.

Inadequate use of certain types of testing contributes to the great disparities in health care services in the U.S.[105] For example, living in a low socioeconomic area has been a key determinant in late-stage diagnosis of colorectal cancer, more so than age, race, gender, and source of care. Poor African-Americans are at increased risk for both occurrence of and mortality from colorectal cancer.[106, 107] One study found that 50% of excess mortality observed in African-Americans was due to late-stage diagnosis.[108] A prominent barrier to adequate screening is lack of health insurance. Providers may not make recommendations for screening or diagnostic tests due to patients' inability to pay for the test or follow-up treatment or providers may focus "triage" to the patients' most pressing health need at the moment.[105] Low-income populations also may be challenged by lack of health literacy, language barriers, negative perceptions about certain diseases, and practical considerations (e.g., job demands).

Multiple reasons are offered regarding why clinicians are not adequately promoting the benefits of well-established screening and preventive services consistent with clinical practice guidelines. Box 11 lists some of the main reasons cited in the literature.

Box 11
Provider Barriers to Preventive Care

- Insufficient time during the clinical encounter

- Organizational barriers (e.g., offering only those preventive services associated with performance measures, administrative costs)

- Provider attitude toward patients (e.g., anticipated lack of patient cooperation with recommended test, inattentiveness to language barriers)

- Provider beliefs and lack of knowledge concerning the evidence supporting the guidelines (e.g., that evidence is inconclusive, insufficient, or conflicting; high number of false negatives)

- Lack of financial incentives encouraging prevention (e.g., capitation, low payment, lack of short-term savings).[21, 105, 109-112]

---

2. *Insufficient Patient Knowledge of High Value Tests*

Patients often expect to receive a laboratory test during physician office visits.[113] They may even see such tests as evidence of sound medical treatment or acknowledgement that their concerns have been heard. Generally, patient expectations are tied to routine screening tests, including blood counts, urine tests, blood chemistries, Pap smears for women, and prostate specific antigen (PSA) tests for men.[114, 115] Patient satisfaction does not seem to be linked to whether expectations were met or not for tests or specialist referrals, although less satisfaction has been reported when new medications are expected but not received.[116]

There are reported marked differences in patient knowledge and requests of certain widely used screening and diagnostic tests, such as FOBT or fecal immunochemical tests (FIT) for colorectal cancer, C-reactive protein for heart disease, and serologic tests for sexually transmitted diseases. For example, a study published in 2005 assessed barriers to screening (via FOBT, FIT, or colorectal endoscopy) in adults 50 years or older who had never been tested or were not current for recommended screening for colon cancer. Patient-related lack of knowledge, awareness, and motivation was cited as the main barrier for about 77% of these individuals and lack of access or provider recommendation were cited as the main barrier for about 22%. Only 10% of adults not current with testing and who had a doctor visit in the past year reported receiving a screening recommendation.[117] Even when patients are aware of screening recommendations, they may not follow through due to other barriers, such as lack of health insurance. Underserved populations tend to be most vulnerable in this regard.[118]

Many providers believe that meeting patient expectations, when medically warranted, is important to satisfaction with care and contributes to health outcomes by improving compliance with treatment and follow-up.[119] However, the actual effect of patient expectations on satisfaction with care is mixed.[113] In some instances, patients may want information rather than specific actions (e.g., test ordering). Patient expectation for testing may vary by the prospects of unfavorable test results or out-of-pocket costs associated with diagnostic tests.[120] Pursuant to patient-centered care, providers must manage patient demands and expectations for laboratory tests and counsel them as to those that would best serve their health needs.[121]

3. Lag Times and Varying Evidence Expectations

Two factors that introduce uncertainty and inconsistency to incorporation of evidence into practice are lag times and varying evidence expectations among organizations that generate clinical practice recommendations. Lag time from appearance of relevant new evidence pertaining to a laboratory test through updating recommendations about use of the test remains a significant challenge to clinicians, patients, payers, and other stakeholders. This applies to any organizations engaged in appraising evidence, including the USPSTF, EGAPP, health professional expert panels, HTA vendors, and others. To the extent that such recommendations are used to inform practice guidelines and coverage policies, this time lag can slow access to beneficial laboratory tests and other interventions. Given its important role in generating evidence-based recommendations pertaining to laboratory testing and other services used in prevention, it is helpful to consider the USPSTF with regard to these factors.

The USPSTF tends to have among the more rigorous evidence requirements of the expert advisory groups that generate recommendations or guidelines for clinical preventive services. Also, the USPSTF is among the organizations that places explicit emphasis on direct evidence of the impact of preventive interventions on patient outcomes, not just intermediate or surrogate outcomes. As such, this can result in "I" statements (insufficient evidence to recommend for or against using a service) rather than "A" or "B" recommendations (to offer or provide a service) in some instances where other organizations recommend offering or providing the service.

Like other public and private sector efforts that generate evidence-based recommendations for clinical practice, the USPSTF is subject to delays from the publication of new evidence on a topic to the time at which this evidence can be incorporated into a new or revised recommendation. Given resource constraints and the backlog of topics, there are time lags between updates of any

topic. As such, relevant evidence may come into the public domain and be incorporated by some expert advisory groups' processes before others.

The USPSTF 2008 update of its 2003 recommendation about screening for prediabetes and diabetes provides a useful example about varying evidence requirements and time lags. In the 2008 update, and drawing on clinical trial evidence generated since the 2003 recommendation, the USPSTF narrowed its grade B recommendation for screening for type 2 diabetes in asymptomatic adults by retaining screening for those with sustained high blood pressure but dropping the recommendation to screen those with hyperlipidemia. Even so, the USPSTF stated that clinicians should assess overall cardiovascular disease risk in patients, and that if the patient's risk is near a threshold for treatment with lipid-lowering drugs, they should screen for diabetes to assess the patient's cardiovascular disease risk. The USPSTF found that there is still insufficient evidence to establish the impact of routine screening for type 2 diabetes on patient outcomes for asymptomatic individuals with normal blood pressure.[122] In 2009, the American Academy of Family Physicians modified its recommendation in a manner that aligned with the USPSTF. In contrast, in its 2009 statement, the American Diabetes Association recommends screening for diabetes and prediabetes beginning at age 45. The Canadian Task Force on Preventive Health Services guideline, last updated in 2005, recommends screening for patients with hypertension or hyperlipidemia. (See accompanying case study on screening for type 2 diabetes and prediabetes.)

In its most recent review of the topic in 2003, the USPSTF concluded that the evidence was insufficient to (1) recommend for or against the routine use of new technologies (i.e., HPV testing, liquid-based cytology, computerized rescreening, and algorithm based screening) as a primary screening tool, or (2) determine whether new technologies are more effective than conventional Pap smear screening, in reducing incidence of or mortality from invasive cervical cancer.[123] Currently, the USPSTF is revisiting the topic of cervical cancer screening and, as such, is expected to consider new evidence from several RCTs published since 2002 confirming the value of adding HPV testing to cervical cytology. (See accompanying case study on HPV screening.)

In screening for hepatitis C virus (HCV), the CDC recommends screening of a broad range of high-risk individuals.[124] In contrast, in its 2004 statement, the USPSTF found insufficient evidence to recommend for or against routine screening for HCV infection in adults at high risk for infection. In particular, USPSTF found no evidence that screening for HCV infection in high-risk adults leads to improved long-term health outcomes, although there is good evidence that antiviral therapy improves intermediate outcomes, such as viral loads of HCV.[125]

IV. Effect of Value Assessments on Reimbursement Policy

A considerable set of challenges to the laboratory medicine sector involves payer reimbursement policies that govern coverage decisions, payment rates, and coding of new tests.[8] Public and private sector insurers use independent processes to conduct HTAs and other evidence appraisals that inform their respective reimbursement decisions. The policies of the Centers for Medicare and Medicaid (CMS), the nation's largest health care purchaser, have an especially strong influence on the reimbursement policies of other federal, state, and private sector payers, particularly with regard to payment rates.[11, 126]

Specific challenges involving payers include the following.

- Coverage of certain laboratory tests based on medical necessity criteria is variable across payer groups, particularly for certain screening tests and for genetic and other newly developed molecular tests.

- Among public and private sector insurers, the multitude of different payment schedules for inpatient and ambulatory care services used through local payer entities can be burdensome to manage and difficult to assess for periodic fee increases, and can result in payment shortfalls and inconsistencies across carriers.

- The standardized coding systems used to list a test on a fee schedule and process claims, as well as the mechanism for updating codes, are inadequate, which leaves providers to use existing codes and underdeveloped code modifiers that lack specificity for newer and emerging tests.

Many of these challenges have been brought to the attention of policy makers in other reports;[8] however, progress in resolving them remains slow.

A. Coverage Decisions

Medicare Coverage. Medicare's authorizing legislation in 1965 established broad categories of coverage for hospital, physician, and laboratory services, but limited payment to expenses deemed reasonable and necessary for the "*diagnosis or treatment* of illness or injury or to improve the functioning of a malformed body member" (emphasis added).[127] This means that the original Medicare statute effectively restricts payment for preventive and screening interventions in patients without signs, symptoms, complaints or personal history of disease or injury, unless otherwise specified by Congress.

In general, policymakers understand the importance of screening and preventive services in improving patient health and reducing expenditures. In 2003 and 2005, Congress expanded Medicare coverage to include certain screening tests, such as those for cardiovascular disease and diabetes.[128] More recently, under the Medicare Improvement for Patients and Providers Act of 2008, CMS gained authority to consider adding preventive services that "that identify medical conditions or risk factors" that the Secretary of Health and Human Services deems "reasonable and necessary for the prevention or early detection of an illness or disability," and those recommended by the USPSTF.[129] As such, the advantages and disadvantages of the USPSTF's approach to developing recommendations for clinical preventive services may be reflected in Medicare coverage policies.

As is so for other payers, Medicare coverage decisions are subject to time lags, reflecting processes to review and revise coverage decisions and the time lag of peer-review publications. Such time lags may delay coverage determinations concerning some innovative technologies and testing methods. Coverage decisions are also affected by the quality of relevant available evidence and some requests for further evidence. The CMS Medicare Evidence Development and Coverage Advisory Committee (MEDCAC) examines available evidence, including in the form of HTAs prepared by EPCs at the request of CMS, in advising CMS regarding pending national coverage determinations (NCDs) for Medicare. For 85% of the various technologies evaluated from 1999 to 2007, including some diagnostic tests and procedures, CMS considered the evidence fair or poor.

Positive NCDs were made in 60% of cases, although almost always with conditions placed on the population or setting to which coverage applies. The conditions for coverage vary, although they increasingly involve restrictions tied to disease severity and treatment regimens as well as requirements for clinical trials, registries, or other data collection.[130]

Consistent with the Medicare statute, most genetic tests are not covered by Medicare unless they are performed on symptomatic individuals or are used to identify treatment-responsive populations.[8, 131] In 2009, CMS[i] is reviewing scientific evidence on genetic tests and assess their value for coverage decisions, including innovative oncology tests and pharmacogenomic tests.

Private Payer Coverage. Private payers maintain their own processes for making coverage decisions and may choose to adjust policies following the introduction of new technologies to the market. In addition to matters of clinical benefit for their beneficiaries, these policies may be affected by a variety of sources, such as state mandates, consumer preference, or financial considerations. Payers often negotiate specific coverage policies with the groups or employers purchasing the health plan.[133] Private payers increasingly draw on HTAs prepared by HTA vendors and other HTA programs and agencies to inform their coverage decisions. Generally, there is consistency in coverage among private sector payers in routine laboratory screening and diagnostic testing associated with standard of care, although some variations remain.[14]

Due in part to how recently these tests have become available and their innovative nature, private payers vary widely in their coverage of pharmacogenomic and other genetic and molecular-based tests, making their own decisions on a case-by-case basis. Consistent with higher and more explicit evidence requirements, payers seek evidence on clinical utility as well as analytical validity and clinical validity for these tests. Private plans generally cover a pharmacogenomic test when it is recommended on a drug label, such as the HER-2/*neu* test to determine whether Herceptin treatment should be recommended for patients with breast cancer.[134] Coverage decisions may vary significantly across individual payers in instances where there evidence is limited on clinical utility and costs associated with a test.[14, 133-135]

B. Payment Systems

Methods of paying for laboratory tests are complex and vary by provider site (e.g., inpatient, outpatient, ambulatory care) and type of test (e.g., clinical or anatomic pathology).[8] In most instances, the Medicare prospective payment systems for inpatient[j] and ambulatory care serve as the basis for federal and private sector payment, although other payers may assign their own payment rates. The ambulatory care payment schemes have proved to be the most challenging for the diagnostics industry.

---

[i] Aside from a few CMS Medicaid-mandated newborn genetic screening tests, coverage decisions are the responsibility of the state Medicaid programs and can vary substantially from state to state.[132]

[j] For inpatient care, technical fees for laboratory services are bundled into diagnosis-related groups (DRGs) based on the patient's diagnosis; ambulatory patient classifications (APC) are used for payment of hospital outpatient surgery.

Laboratory tests for ambulatory care are paid according to predetermined, fixed-fee schedules, negotiated contracts, or competitive bidding contracts.[k] The main CMS fee schedules are the:

- Medicare physician fee schedule (MPFS), which covers clinician services, including pathologist interpretive services for certain anatomic pathology or molecular tests

- Medicare Clinical Laboratory Fee Schedule (CLFS) for payment of clinical laboratory tests

Private payers often use the MPFS and CLFS for setting their own payment rates as a multiple or percentage of Medicare rates, although private payers typically negotiate contracts with providers and suppliers, including clinical laboratories.

Policies related to the CLFS have proven to be the greatest challenge to laboratories for several reasons. The CLFS was established in 1984 and is based on prevailing charges from 1983. The CLFS contains payment ceilings, or national limitation amounts (NLA). While Congress originally intended to adjust the CLFS annually for inflation, this has rarely occurred.[14] Over the past 25 years, the NLAs have been reduced seven times; also, there were only two pricing updates, 1.1% (2003) and 4.5% (2009), during the period 1997-2009.[11, 136] Because the CLFS historically has been managed at the local level by 56 local carriers, there have been essentially 56 different fee schedules until recently. Now that CMS has consolidated these carriers into 15 Medicare Administrative Contractor (MAC) jurisdictions, some consolidation of these fee schedules is occurring. According to a recent report by the Office of the Inspector General (OIG) at HHS, 83% of carrier rates were at the NLAs and 89% of claims are paid at the NLAs.[137] However, 97% of lab tests had at least one carrier rate that varied from the NLA. For some of the most frequently used tests, OIG found wide variations in payment levels, including variations of up to 40% for complete blood count, 30% for a urinalysis, 28% for glycosylated hemoglobin, 27% for lipid panel, 19% for prothrombin time, and 11% for basic metabolic panel.[137] As noted in our earlier report and elsewhere, the CMS processes for developing CLFS payment rates for new test HCPCS codes pose significant uncertainties and other challenges for laboratories and manufacturers that hinder incentives to innovate in this area.[8, 94]

V. Key Findings and Conclusions

Overarching conclusion:

Innovation, demonstrated clinical benefit, and appropriate use of laboratory screening and diagnostic tests are essential to achieving goals of health system reform. Clinical laboratory testing is integral to evidence-based improvements in health care quality, patient outcomes, efficiency, and accountability.

1. ***Screening and diagnostic tests contribute to health care value across the spectrum of care.*** Information from clinically appropriate testing contributes to early detection, diagnosis, patient and clinician decision-making, choice of treatment, therapeutic monitoring, reducing adverse events, improved health outcomes and quality of life, and more cost-effective care.

---

[k] Refer to the CDC-sponsored report, *Laboratory Medicine: A National Status Report*, for a full review of reimbursement policy and payment calculation methodology at http://www.futurelabmedicine.org

2. *Laboratory testing, including existing, new, and emerging testing technologies, aligns with and will have an integral role in meeting major goals of national health reform.* These aspects include wider access to testing; early detection and treatment; support of personalized medicine and other tailored services for priority populations and others; more appropriate, efficient and cost-effective care for chronic conditions; slower disease progression; and faster recovery and less disability.

3. *A growing body of evidence demonstrates the value of laboratory screening tests in primary prevention*—that is, in early detection at a time when it is possible to prevent the onset of disease, not just treat disease after it has occurred. Primary prevention offers opportunities for cost-effective care and net savings in some instances. Advances in genetic and molecular testing are enhancing the potential of primary prevention through identification of individuals with pre-dispositions for disease.

4. *Key HTA and clinical practice guideline groups are setting high evidence bars, emphasizing explicitly the need for demonstrating the clinical utility of tests used in screening and diagnosis.* In addition to analytic validity (test accuracy, precision, robustness), which is typically established by regulatory processes, and clinical validity (detect and predict probability of having a disorder based on a test result), there is greater demand for evidence of clinical utility (impact on clinical outcomes and usefulness to patient and clinician decision-making) by the USPSTF, EGAPP, various HTA agencies, private payers, and others.

   - RCTs that provide direct evidence of impact of testing on patient outcomes remain preferred to other study designs and indirect evidence of such impact. This remains a challenge to, and can be impractical for, demonstrating value of testing, particularly given the confounding effects of intervening decisions, interventions, and environmental factors between testing and ultimate outcomes, and the costs and time (which can be years or decades) needed to assess impact on outcomes using RCTs.

   - Evidence hierarchies and grading systems, such as recently revised ones by the USPSTF and those of EGAPP and GRADE, should continue to be adapted for screening and diagnostic testing, accounting for emerging, practical study designs for validating tests and accompanying evidence appraisal methods.

   - The explicitness of evidence expectations by regulators, guideline groups, and other gatekeepers provides opportunities for test manufacturers and clinical laboratories to build evidence requirements into innovation, development, validation, and marketing of laboratory tests.

5. *Varying evidence standards and time lags for assessing evidence and updating recommendations and guidelines introduce inconsistency and uncertainty to incorporating evidence about some laboratory tests into practice.* Lag time from appearance of relevant new evidence through updating and disseminating recommendations can affect health care quality and access. Different evidence requirements, including the generally more rigorous ones of USPSTF compared to others, along with varying review cycles of these groups, can result in divergent findings in such areas as screening for diabetes, human papillomavirus, and hepatitis C virus. These factors also pose risk and uncertainty to laboratories and test manufacturers.

6. *New methods and analytical tools are emerging for assessing and demonstrating the clinical and economic impact of laboratory tests*. While RCTs remain the preferred study design for establishing causal effects of interventions on patient outcomes, strengthening and emergence of other study designs are helping to supplement and, in some instances, may substitute for traditional RCTs. Included are variations in traditional clinical trial designs; "data mining" of claims data, patient registries, and electronic medical records; retrospective studies of specimen remnants; and analyses of linked data sets of laboratory data and patient outcomes.

7. *Laboratory testing has prominent roles in the national agenda for comparative effectiveness research that provide opportunities for broad demonstration of value in "real-world" health care*. Two main roles for laboratory testing in CER are: (1) for patient selection and tracking intermediate and long-term health outcomes in CER of other interventions, and (2) as the index interventions for CER, e.g., in head-to-head comparisons of alternative laboratory tests or comparisons of laboratory tests to other tests for particular health care conditions. Included among the Institute of Medicine's recommended top national CER priorities are screening for methicillin resistant *S. aureus* (MRSA) and genetic and biomarker testing for multiple types of cancer.

8. *Greater investment is needed in the small but growing body of evidence on the economic impact of laboratory testing*. As is so for many other types of health care technology, wider demand by clinicians, provider institutions, payers and policy makers, including in national deliberations on health reform, calls for evidence pertaining to the distinct analyses of cost-savings, cost-effectiveness, and cost-utility of laboratory testing. Findings about laboratory tests that are truly cost-saving or highly cost-effective, such as colorectal cancer screening in adults aged 50-75 and screening young women for chlamydial infection, should be broadly recognized and used as models for further work. Key considerations are that economic impact of specific tests can vary significantly depending on the population targeted (with testing of high-risk populations more likely to be cost-effective) and tradeoffs between the cost of greater testing frequency and yield of cases detected.

9. *Laboratory testing has a central role in personalized medicine, whose extraordinary potential is recently emerging into practice*. Recent scientific and technological advances have led to molecular-level and genetic testing, including pharmacogenomics, that enable tailoring therapies to subgroups and individuals, i.e., to ensure 'the right treatment for the right patient at right time.' In parallel, EGAPP and other groups are developing evidence frameworks to guide assessments of such genetic and genomic testing technologies.

10. *Payment policies that govern coverage decisions, payment rates, and coding of new tests remain major challenges to the laboratory medicine sector*. Coverage of certain laboratory tests, particularly certain screening tests and for genetic and other newly developed molecular tests, is variable across payers. The multitude of payment schedules for inpatient and ambulatory care services used through local payer entities can be burdensome to manage and difficult to assess for periodic fee increases, and can result in payment shortfalls. The coding systems used to list tests on fee schedules and process claims, and the mechanisms to update them, are inadequate, often leaving providers to use existing codes and underdeveloped code modifiers that are not specific to newer and

emerging tests. Medicare's processes for developing CLFS payment rates for new test HCPCS codes pose significant uncertainties and other challenges for laboratories and manufacturers that hinder the incentive to innovate in this area. Progress in resolving these longstanding problems remains slow.

References

1. Reinhardt UE, Hussey PS, Anderson GF. US health care spending in an international context. Health Aff (Millwood);25(1):186-96.

2. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348(26):2635-45.

3. Mangione-Smith R, DeCristofaro AH, Setodji CM, et al. The quality of ambulatory care delivered to children in the United States. N Engl J Med 2007;357:1515-23.

4. Institute of Medicine. Building a better delivery system: a new engineering/health care partnership. Washington, DC: National Academy Press, 2005.

5. Orszag PR. Long-term budget outlook and options for slowing the growth of health care costs. Washington, DC: Congressional Budget Office, 2008.

6. Porter ME. What is value in health care? Presentation at 2007 Institute of Medicine Annual Meeting: Evidence-based medicine and the changing nature of health care, October 8, 2007.

7. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academy Press, 2001.

8. Laboratory medicine: a national status report. The Lewin Group, 2008. (Accessed July 5, 2009, at http://www.futurelabmedicine.org.)

9. Price CP. Evidence-based laboratory medicine: supporting decision-making. Clin Chem 2000;46(8 Pt 1):1041-50.

10. OMB historical tables. Table 16.1--Outlays for health programs: 1962-2014. Office of Management and Budget, 2009. (Accessed September 22, 2009, at http://www.whitehouse.gov/omb/budget/Historicals/.)

11. Terry M. Lab industry strategic outlook: market trends and analysis 2007. New York, NY: Washington G-2 Reports, 2007.

12. National health expenditures by type of service and source of funds, CY 1960-2007. Centers for Medicare and Medicaid Services, 2009. (Accessed September 22, 2009, at http://www.cms.hhs.gov/NationalHealthExpendData/02_NationalHealthAccountsHistorical.asp#TopOfPage.)

13. Price CP. Application of the principles of evidence-based medicine to laboratory medicine. Clinica Chimica Acta 2003;333:147-54.

14. The Lewin Group. The value of diagnostics innovation, adoption and diffusion into health care. Advanced Medical Technology Association, 2005. (Accessed July 19, 2009, at http://www.advamed.org/NR/rdonlyres/61EB858F-EC9E-4FAB-9547-09DABF7D2A72/0/thevalueofdiagnostics.pdf.)

15. Value of clinical laboratory services in health care. Bethesda, MD: American Society for Clinical Laboratory Science, 2005. (Accessed May 15, 2007, at http://www.ascls.org/position/ValueCLS.asp.)

16. GeneTests. Seattle, WA: GeneTests.org, 2008. (Accessed April 18, 2008, at http://www.genetests.org/servlet/access?id=8888891&key=Z1o3AV3PRzDMu&fcn=y &fw=apWV&filename=/.)

17. Hill C. Automating nucleic acid amplification tests. IVD Technology 2000.

18. Csako G. Present and future of rapid and/or high-throughput methods for nucleic acid testing. Clin Chim Acta 2006;363(1-2):6-31.

19. Master list of references and glossary of genetic terms. ACCE Model System for Collecting, Analyzing and Disseminating Information on Genetic Tests, 2009. (Accessed July 20, 2009, at http://www.cdc.gov/genomics/gtesting/file/print/FBR/CF_Glossary.pdf.)

20. A purchaser's guide to clinical preventive services: moving science into coverage. Washington, DC: National Business Group on Health, 2006.

21. Amonkar MM, Madhavan S, Rosenbluth SA, simon KJ. Barriers and facilitators to providing common preventive screening services in managed care settings. J Community Health 1999;24(3):229-47.

22. Maciosek MV, Coffield AB, Edwards NM, Flottemesch TJ, Goodman MJ, Solberg LI. Priorities among effective clinical preventive services: results of a systematic review and analysis. Am J Prev Med 2006;31(1):52-61.

23. Health, United States, 2007. Hyattsville: National Center for Health Statistics, 2008.

24. Ransohoff D. Colon cancer screening in 2005: status and challenges. Gastroenterology 2005;128:1685-95.

25. Guide to clinical preventive services, 2007. Recommendations of the U.S. Preventive Services Task Force. U.S. Preventive Services Task Force, 2007. (Accessed July 19, 2009, at www.ahrq.gov/clinic/pocketgd.htm.)

26. Medicare National Coverage Determinations (NCD) coding policy manual and change report, January 2009. Baltimore: Centers for Medicare and Medicaid,

27. Goodman CS. HTA 101: introduction to health technology assessment. National Institutes of Health, 2004. (Accessed July 15, 2009, at http://www.nlm.nih.gov/nichsr/hta101/hta101.pdf.)

28. Institute of Medicine. Assessing medical technologies. Washington, DC: National Academy Press, 1985.

29. UK National Health Service R&D Health Technology Assessment Programme. National Health Service, 2003. (Accessed July 21, 2009, at http://www.hta.nhsweb.nhs.uk/abouthta.htm.)

30. Institute of Medicine. Performance measurement: accelerating improvement. Washington, DC: National Academy Press, 2006.

31. Institute of Medicine. Valuing health for regulatory cost-effectiveness analysis. Washington DC: The National Academies Press, 2006.

32. Ward-Cook KM, Lehmann CA, Schoeff LE, Williams RH. Clinical diagnostic technology: the total testing process. Volume 2: the analytical phase. Washington, DC: American Association of Clinical Chemistry, 2005.

33. Steurer J, Fischer JE, Bachmann LM, Koller M, ter Riet G. Communicating accuracy of tests to general practitioners: a controlled study. Br Med J 2002;324:824-6.

34. Tape TG. Interpreting diagnostic tests. University of Nebraska Medical Center, 2008. (Accessed July 4, 2008, at http://gim.unmc.edu/dxtests/Default.htm.)

35. Poon LM, Chan KH, Smith GJ, et al. Molecular detection of a novel human influenza (N1H1) of pandemic potential by conventional and real-time quantitative RT-PCR assays. Clin Chem 2009;0.

36. Foo H ,Dwyer DE. Rapid tests for the diagnosis of influenza. Australian Prescriber, 2009. (Accessed May 21, 2009, at http://www.australianprescriber.com/upload/pdf/news/182.pdf.)

37. Hellems MA, Hayden GF. Rapid diagnostic testing for influenza: when does it make sense? Consultant for Pediatrician 2006;5(12):1-7.

38. Genetic testing for fragile X. National Fragile X Foundation, 2009. (Accessed May 22, 2009, at http://www.fragilex.org/html/testing.htm.)

39. Zhou Y, Lum J, Yeo GH, et al. Simplified molecular diagnosis of fragile X syndrome by fluorescent methylation-specific PCR and GeneScan analysis. Clin Chem 2006;52:1492-500.

40. Ng K, Zhu AX. Targeting the epidermal growth factor receptor in metastatic colorectal cancer. Crit Rev Oncol Hematol 2008;65(1):8-20.

41. Brown HM, Aronson N, Ziegler KM, et al. KRAS mutations and epidermal growth factor receptor inhibitor therapy in metastatic colorectal cancer. Report No. 23(6). Technology Evaluation Center, Blue Cross Blue Shield, 2009.

42. Van Cutsem E, Lang I, D'haens G, et al. KRAS status and efficacy in the first-line treatment of patients with metastatic colorectal cancer (mCRC) treated with FOLFIRI with or without cetuximab: the CRYSTAL experience. J Clin Oncol 2008;26(Supplement):2.

43. Bokemeyer I, Bondarenko JT, Hartmann FG, et al. KRAS status and efficacy of first-line treatment of patients with metastatic colorectal cancer (mCRC) with FOLFOX with or without cetuximab: the OPUS experience. J Clin Oncol 2008;26(Supplement):4000.

44. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med 1978;299:926-30.

45. Heim SW, Schectman JM, siadaty MS, Philbrick JT`. D-dimer testing for deep vein thrombosis: a meta-analysis. Clin Chem 2004;50(7):1136-47.

46. Rückerl R, Peters A, Khuseyinova N, et al. Determinants of the acute-phase protein C-reactive protein in myocardial infarction survivors: the role of comorbidities and environmental factors. Clin Chem 2009;55:322-35.

47. Leeflang M, Reitsma J, Scholten R. Impact of adjustment for quality on results of meta-analyses of diagnostic accuracy. Clin Chem 2007;53:164-72.

48. U.S. system of oversight of genetic testing: a response to the charge of the Secretary of Health and Human Services. Bethesda: Secretary's Advisory Committee on Genetics, Health, and Society, 2008.

49. Whitlock EP, Lin J, Liles E, et al. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. Ann Intern Med;149(9):638-58.

50. Hewitson P, Glasziou P, Irwig L, et al. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. The Cochrane Collaboration, 2007.

51. Mandel JS, Church TR, Ederer F, et al. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. J Natl Cancer Inst 1999;91:434-7.

52. Screening for colorectal cancer. Recommendation statement. U.S. Preventive Services Task Force, 2008. (Accessed August 20, 2009, at http://www.ahrq.gov/clinic/uspstf08/colocancer/colors.htm.)

53. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001;69:-89.

54. Institute of Medicine. Initial national priorities for comparative effectiveness research. Washington, DC: The National Academies, 2009.

55. AlloMap HTx Molecular Expression Testing. Laboratory services guide. Brisbase: XDx, 2009.

56. Mayrand MH, Duarte-Franco E, Rodrigues I, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. N Engl J Med 2007;357(16):1579-88.

57. Mennemeyer ST, Winkelman JW, Kiefe CI. Monitoring patient outcomes after clinical laboratory testing. In: Laboratory-related measures of patient outcomes: an introduction. Bissell MG, ed. Washington, DC: AACC Press, 2000.

58. Bissell MG. Introduction: what's in a laboratory outcome? In: Laboratory-related measures of patient outcomes: an introduction. Bissell MG, ed. Washington, DC: American Association for Clinical Chemistry, 2000.

59. Laboratory medicine best practices: developing an evidence-based recommendation process. Final technical report. Atlanta, GA: Centers for Disease Control and Prevention, 2007.

60. Bruns DE. Laboratory-related outcomes in healthcare. Clin Chem 2001;47(8):1547-52.

61. Birkmeyer JD, Ker EA, Dimick JB. Improving the quality of quality measurement. In: Performance measurement: accelerating improvement. Washington, DC: National Academy Press, 2006.

62. Beck JR. Pathologist as a designer of process improvement. CAP Foundation Conference VIII. Patient-centered pathology practice: outcomes and accountability. February 2, 1997. St. Petersburg, FL: College of American Pathologists, 1997.

63. Bissell MG, Jones H. The laboratory's role in outcome assessment. MLO Med Lab Obs 1998;30(1):54, 56-4, 58.

64. Budgetary effects of proposals to expand governmental support for preventive medical care and wellness services. Letter to Hon. Nathan Deal. Congressional Budget Office, 2009. (Accessed August 15, 2009, at http://www.cbo.gov/ftpdocs/104xx/doc10492/08-07-Prevention.pdf.)

65. Scott RD. Direct medical costs of healthcare-associated infections in U.S. hospitals and the benefits of prevention. Atlanta: Centers for Disease Control and Prevention, 2009.

66. Raygada JL, Levine DP. Methicillin-resistant *Staphylococcus aureus*: a growing risk in the hospital and in the community. American Health and Drug Benefits 2009;2(2):86-95.

67. Boyce JM, Jackson MM, Pugliese G, et al. Methicillin-resistant *Staphylococcus aureus* (MRSA): a briefing for acute care hospitals and nursing facilities. The AHA Technical Panel on Infections Within Hospitals. Infection Control and Hospital Epidemiology 1994;15:105-15.

68. Huang SS, Platt R. Risk of methicillin-resistant *Staphylococcus aureus* infection after previous infection or colonization. Clin Infect Dis 2003;43:971-8.

69. VHA Directive 2007-002. Methicillin-resistant *Staphylococcus aureus* (MRSA) initiative. Veterans Health Administration, 2007.

70. Felder RA. Robotics and automated workstations for rapid response testing. Am J Clin Pathol 1995;104(4 Suppl 1):S26-S32.

71. De Cresce RP, Phillips DL, Howanitz PJ. Financial justification of alternate site testing. Arch Pathol Lab Med 1995;119(10):898-901.

72. Howanitz PJ, Jones BA. Comparative analytical costs of central laboratory glucose and bedside glucose testing: a College of American Pathologists Q-Probes study. Arch Pathol Lab Med 2004;128(7):739-45.

73. Haddix AC, Teutsch SM, Corso PS. Prevention effectiveness: a guide to decision analysis and economic evaluation. Oxford, United Kingdom: Oxford University Press, 2003.

74. Cost-effectiveness in health and medicine. Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. New York, NY: Oxford University Press, 1996.

75. Veenstra DL, Higashi MK, Phillips KA. Assessing the cost-effectiveness of pharmacogenomics. AAPS PharmSci 2000;2(3):1-11.

76. Carlson JJ, Henrikson NB, Veenstra DL, Ramsey SD. Economic analyses of human genetics services: a systematic review. Genet Med 2005;7(8):519-23.

77. Hu D, Hook EW, III, Goldie SJ. Screening for *Chlamydia trachomatis* in women 15 to 29 years of age: a cost-effectiveness analysis. Ann Intern Med 2004;141(7):501-13.

78. Decision memo for screening immunoassay fecal-occult blood test (CAG-00180N). Baltimore, MD: Centers for Medicare and Medicaid Services, 2007. (Accessed June 10, 2007, at http://www.cms.hhs.gov/mcd/viewdecisionmemo.asp?id=87.)

79. Dendukuri N, Khetani K, McIsaac M, Brophy J. Testing for HER2-positive breast cancer: a systematic review and cost-effectiveness analysis. CMAJ 2007;176(10):1429-34.

80. Marshall DA, Kleinman SH, Wong JB, et al. Cost-effectiveness of nucleic acid test screening of volunteer blood donations for hepatitis B, hepatitis C and human immunodeficiency virus in the United States. Vox Sang 2004;86(1):28-40.

81. Paltiel AD, Weinstein MC, Kimmel AD, et al. Expanded screening for HIV in the United States: an analysis of cost-effectiveness. N Engl J Med 2005;352(6):586-95.

82. Wu O, Robertson L, Twaddle S, et al. Screening for thrombophilia in high-risk situations: systematic review and cost-effectiveness analysis. The Thrombosis: Risk and Economic assessment of Thrombophilia Screening (TREATS) study. London: Health Technology Assessment Program, National Health Service, 2006.

83. Cohen JT, Neumann PJ, Weinstein MC. Does preventive care save money? Health economics and the presidential candidates. N Engl J Med 2008;358(7):661-3.

84. Harris RP, Helfand M, Woolf SH, et al. Current methods of the U.S. Preventive Services Task Force: a review of the process. Am J Prev Med 2001;20(3S):21-35.

85. Sawaya GF, Guirguis-Blake J, LeFevre, et al. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. Ann Intern Med 2007;147(12):871-5.

86. Teutsch SM, Bradley LA, Palomaki GE, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) initiative: methods of the EGAPP Working Group. Genet Med 2009;11(1):3-14.

87. Technology Evaluation Center assessments. Chicago, IL: Blue Cross Blue Shield Association, 2007. (Accessed November 26, 2007, at http://www.bluecares.com/betterknowledge/tec/tec-assessments.html#topics.)

88. Melanson SE, Lindeman NJ, Jarolim P. Selecting automation for the clinical chemistry laboratory. Arch Pathol Lab Med 2007;131:1063-9.

89. U.S. Preventive Services Task Force procedure manual. U.S. Preventive Services Task Force, 2008. (Accessed June 20, 2009, at http://www.ahrq.gov/Clinic/uspstf08/methods/procmanual.htm.)

90. Schünemann HJ, Oxman AD, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. Br Med J 2008;336(7653):1106-10.

91. Ebell MH, Siwek J, Weiss BD, et al. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. Am Fam Physician 2004;69(3):548-56.

92. Methods of reference guide for effectiveness and comparative effectiveness reviews. Gaithersburg: Agency for Healthcare Research and Quality, 2007.

93. Guyatt GH, Tugwell PX, Fenney DH. A framework for clinical evaluation of diagnostic technologies. Can Med Assoc J 1986;134:587-94.

94. Sawaya GF, Sung HY, Kinney W, Kearney KA, Miller MG, Hiatt RA. Cervical cancer after multiple negative cytologic tests in long-term members of a prepaid health plan. Acta Cytol 2005;49(4):391-7.

95. Furberg BD, Furberg CD. What are the strengths of observational studies? In: Evaluating clinical research. New York: Springer, 2007.

96. Institute of Medicine Roundtable on Evidence-based Medicine. The learning health system. Washington, DC: The National Academies, 2004.

97. McGlynn EA, Cassel CK, Leatherman ST, DeCristofaro A, Smits HL. Establishing national goals for quality improvement. Med Care 2003;41(1 supplement):I16-I29.

98. Walsh JM, McPhee SJ. Prevention in the year 2002: some news, some issues. Primary Care 2002;29:727-49.

99. American Cancer Society guidelines for the early detection of cancer. American Cancer Society, 2009. (Accessed July 23, 2009, at http://www.cancer.org/docroot/ped/content/ped_2_3x_acs_cancer_detection_guidelines_36.asp.)

100. 2006 National Healthcare Quality Report. Report No. 07-0013. Rockville, MD: Agency for Healthcare Research and Quality, 2006.

101. Sirovich BE, Welch HG. Cervical cancer screening among women without a cervix. JAMA 2004;291:2990-3.

102. Schober SE, Carroll MD, Lacher DA, Hirsch R. High serum total cholesterol--an indicator for monitoring cholesterol lowering efforts: U.S. adults, 2005-2006. Hyattsville: National Center for Health Statistics, 2007.

103. 2004 National Healthcare Quality Report. Gaithersburg: Agency for Healthcare Research and Quality, 2004.

104. Grant RW, Buse JB, Meigs JB. Quality of diabetes care in U.S. academic medical centers. Diabetes Care 2005;28:337-442.

105. O'Malley AS, Beaton E, Yabroff KR, Abramson R, Mandelblatt J. Patient and provider barriers to colorectal cancer screening in the primary care safety-net. Prev Med 2004;39:56-63.

106. Ries LA, Wingo PA, Miller DS, et al. The annual report to the nation on the status of cancer 1973-1997, with a special section on colorectal cancer. Cancer 2000;88:2398-424.

107. Mandelblatt J, Andrews H, Kao R, Wallace R, Kerner J. The late-stage diagnosis of colorectal cancer: demographic and socioeconomic factors. Am J Public Health 1996;86(1794):1797.

108. Mayberry RM, Coates RJ, Hill HA, et al. Determinants of black/white differences in colon cancer survival. J Natl Cancer Inst 1995;87:1686-93.

109. Wee CC, McCarthy EP, Phillips RS. Factors associated with colon cancer screening: the role of patient factors and physician counseling. Prev Med 2005;41:23-9.

110. Hulscher ME, Wensing M, Grol RP, van der Weijden T, van Weel C. Interventions to improve the delivery of preventive services in primary care. Am J Public Health 1999;89:737-46.

111. Anderson LA, Janes GR, Jenkins C. Implementing preventive services: to what extent can we change provider performance in ambulatory care? A review of the screening, immunization, and counseling literature. Ann Behav Med 1998;20:161-7.

112. Stone EG, Mortob SC, Hulscher ME, et al. Interventions that increase the use of adult immunization and cancer screening services. Ann Intern Med 2002;136:641-51.

113. Rao JK, Weinberger M, Kroenke K. Visit-specific expectations and patient-centered outcomes. Arch Fam Med 2000;9:1148-55.

114. Froehlich GW, Welch HG. Meeting walk-in patients' expectations for testing. Effects on satisfaction. J Gen Intern Med 1996;11:470-4.

115. Kravitz RL, Callahan EJ, Azari R, Antonius D, Lewin CE. Assessing patient' expectatons in ambulatory medical practice. Does the measurement approach make a difference? J Gen Intern Med 1997;12:67-72.

116. Peck BM, Ubel PA, Roter DL, et al. Do unmet expectations for specific tests, referrals, and new medications reduce patients' satisfaction? J Gen Intern Med 2004;19(11):1080-7.

117. Klabunde CN, Vernon SW, Nadel MR, Breen N, Seeff LC, Brown ML. Barriers to colorectal cancer screening: a comparison of reports from primary care physicians and average-risk adults. Med Care 2005;43(9):939-44.

118. O'Malley AS, Forrest CB, Mandelblatt J. Adherence of low-income women to cancer screening recommendations. The roles of primary care, health insurance, and HMOs. J Gen Intern Med 2002;17:144-54.

119. Ware JE, Davies AR. Behavioral consequences of consumer dissatisfaction with medical care. Eval Program Plann 1983;6:291-7.

120. Woolf SH, Kamerow DB. Testing for uncommon conditions: the heroic search for positive test results. Arch Intern Med 1990;150:2451-8.

121. Kravitz RL, Cope DW, Bhrany V, Leake B. Internal medicine patients' expectations for care during office visits. J Gen Intern Med 1994;9:75-81.

122. U.S. Preventive Services Task Force. Screening for type 2 diabetes mellitus in adults: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2008;148(11):846-54.

123. Screening for cervical cancer. Recommendations and rationale. U.S. Preventive Services Task Force, 2003. (Accessed August 20, 2009, at http://www.ahrq.gov/clinic/3rduspstf/cervcan/cervcanrr.pdf.)

124. Hepatitis C. Centers for Disease Control and Prevention, 2008. (Accessed May 23, 2009, at http://www.cdc.gov/hepatitis/HCV.htm.)

125. Screening for hepatitis C in adults. U.S. Preventive Services Task Force, 2004. (Accessed May 30, 2009, at http://www.ahrq.gov/clinic/uspstf/uspshepc.htm.)

126. Institute of Medicine. Medicare laboratory payment policy: now and in the future. Washington, DC: National Academy Press, 2000.

127. Neumann PJ, Divi N, Beinfeld MT, et al. Medicare's national coverage decisions, 1999-2003: quality of evidence and review times. Health Aff (Millwood) 2005;24(1):243-54.

128. Summary of H.R. 1, Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Public Law 108-173. Centers for Medicare and Medicaid, 2004. (Accessed October 1, 2007, at http://www.cms.hhs.gov/MMAUpdate/downloads/PL108-173summary.pdf.)

129. H.R. 6331. Medicare Improvement for Patients and Providers Act of 2008. GovTrack.us, 2009. (Accessed July 20, 2009, at http://www.govtrack.us/congress/bill.xpd?bill=h110-6331.)

130. Neumann PJ, Kamae MS, Palmer JA. Medicare's national coverage decisions for technologies, 1999-2007. Health Aff (Millwood) 2008;27(6):1620-31.

131. Realizing the promise of pharmacogenomics: opportunities and challenges. Draft report of the Secretary's Advisory Committee on Genetics, Health, and Society. Bethesda, MD: Secretary's Advisory Committee on Genetics, Health, and Society, 2007.

132. Newborn screening: characteristics of state programs. Washington DC: Government Accountability Office, 2003.

133. Schoonmaker MM. Reimbursement for molecular diagnostics. IVD Technology 2007.

134. Coverage and reimbursement of genetic tests and services. Secretary's Advisory Committee on Genetics, Health, and Society, 2006. (Accessed October 10, 2006, at http://www4.od.nih.gov/oba/sacghs/reports/CR_report.pdf.)

135. Schoonmaker MM. Private health insurance coverage and payment policies and decision-making processes for genetic technologies and services. SACGHS March 2004 Meeting. March 1, 2004. Bethesda, MD: Secretary's Advisory Committee on Genetics, Health, and Society, 2007.

136. CMS Manual System: Pub 100-04, Medicare claims processing, transmittal 1651, change request 6070. Centers for Medicare and Medicaid Services, 2008. (Accessed July 23, 2009, at http://www.aab.org/Transmittall%201651%202009%20CLFS.pdf.)

137. Office of the Inspector General. Variation in the Clinical Laboratory Fee Schedule. Report No. OEI-05-08-00400. Washington DC: Department of Health and Human Services, 2009. http://www.oig.hhs.gov/oei/reports/oei-05-08-00400.pdf.

Appendix
Case Studies

Case Study I:

Rapid Testing for Methicillin-Resistant Staphylococcus Aureus (MRSA)

Magnitude and Importance of MRSA

Health care-acquired infections (HAIs) are caused by bacteria and other microorganisms that may be resistant to one or more classes of antimicrobial agents. HAIs are a major cause of medical complications and death, and account for increased treatment costs around the world.[1] According to the Centers for Disease Control and Prevention (CDC), nearly 2 million patients suffer from an HAI in U.S. hospitals each year, resulting in 99,000 deaths and $20 to $45 billion in additional health care costs.[2, 3]

The majority of HAIs occur in hospitals, particularly intensive care units, and long-term care facilities, but are transmissible in any health care setting. Bacterial pathogens can be transmitted to patients from contaminated health care workers' hands, clothing, or other devices; contaminated medical equipment (e.g., surgical equipment); catheter insertion and maintenance; and patient-to-patient contact. Opportunities for transmission increase dramatically as patients transfer between health care settings. In many instances, individuals are colonized (i.e., carriers) but do not have signs and symptoms.

Multidrug resistant organisms (MDRO) refer to HAI-related pathogens that are not susceptible to certain antibiotics. The most common MDRO[a] is methicillin-resistant *Staphylococcus aureus*[b] (MRSA)—pathogens resistant to penicillins and cephalosporins. In the early 1990s, MRSA accounted for 20-25% of all *Staphylococcus aureus (S. aureus;* which includes methicillin-susceptible *S. aureus [MSSA])* isolates[c] identified in hospitalized patients. However, by 2007, that proportion increased to 50% based on data reported to the CDC's National Healthcare Safety Network.[5, 6] Other reports indicate that MRSA causes about 30% of surgical site, 24% of ventilator-related pneumonia, 10% of central line, and 2% urinary catheter-associated bloodstream infections. Across all ICUs (e.g., cardiac, pediatric, surgical), 70% of HAIs are MRSA-related.[6]

Evidence is accumulating to justify expanded MRSA screening practices in health care. According to a national survey conducted from 2001 to 2004 (the most recent data available), the prevalence distribution of *S. aureus* colonization is changing. While the overall prevalence of *S. aureus* in patients decreased from 32.4% (2001-2002) to 28.6% (2003-2004), the percentage of those with methicillin-resistant *S. aureus* increased from 0.8% to 1.5% during that same time span.[7] This means that roughly 89 million Americans over age 1 year are undetected carriers of *S. aureus*, of which 1.3 million are MRSA carriers. This subpopulation serves as a reservoir for person-to-person transmission in health care settings. Screening tests for MRSA identify carriers so that appropriate and timely control measures can be implemented to minimize the spread of MRSA to patients who are in weakened, compromised conditions and more likely to acquire a hospital induced infection.

---

[a] Other MDRO pathogens include vancomycin-resistant *Enterococcus* species (VRE), multi-drug resistant gram-negative bacilli, and vancomycin-resistant *Staphylococcus aureus.*

[b] MRSA was first identified in England in 1961, two years after the introduction of methicillin, and first reported in the U.S. in 1968.[4]

[c] An isolate is a pure microbial or viral sample that has been obtained from an infected individual.

New strains of MRSA are evolving rapidly. Community-associated MRSA[d] (CA-MRSA) poses a great public health concern, as certain strains have genetic qualities that allow them to mutate quickly and transmit readily over wide geographic areas, resulting in more severe and atypical manifestations of the disease.[4] As a result, individuals increasingly present to emergency rooms with CA-MRSA (about 60% of emergency room visits for skin and soft tissue infections). Changing resistance patterns and strains of MRSA in ICUs reflect the growing threat of CA-MRSA. Some experts contend that CA-MRSA may overtake health care-acquired MRSA in coming years.[8]

Thus, an active surveillance program to screen patients as they enter the health care system and monitor them regularly during extended stays is the most effective method for detecting underlying pathogens and antibiotic resistance in a growing proportion of the population.[9, 10]

High Economic Burden

The economic burden of HAIs is substantial. The major contributors to direct hospital costs associated with MRSA infections include increased hospital length of stay, ICU length of stay, use of single rooms for isolation, use of more expensive antibiotics to treat MRSA, and use of laboratory tests for monitoring of infections.[11-13]

Adjusted to 2007 dollars using the consumer price index (CPI) for inpatient hospital services, CDC estimates that the direct cost per case for all HAIs ranges from $20,549 to $25,903, for a total of roughly $36 to 45 billion in annual costs.[3] Costs per case of bacteremia also tend to be higher for MRSA than for MSSA. Patients with bacteremia due to MRSA have nearly double the mortality rate, 30% longer hospital stays, and 36% to 40% higher hospital costs than patients with MSSA.[14] This economic burden is accentuated by the indirect costs to patients and society resulting from lost income and productivity.

Data from CDC's National Healthcare Safety Network reported that the most common source of bacteremia was central line catheters, with infection rates of 38% among patients with these catheters.[6] Table 1 provides a summary of the per-patient hospital and total annual hospital costs for all HAIs and selected sites of infection as estimated by CDC.

Table 1. Estimated per patient and total costs for the most frequently occurring device- and procedure-related sites of infection based on 2007 CPI for inpatient hospital services

	No. of infections	Estimated per patient costs	Estimated total annual costs
All HAIs	1,737,125	$20,549 - $25,903	$35.7 - 45 billion
SSI	290,485	$11,874 - $34,670	$3.5 - 10 billion
CLABSI	92,011	$7,288 - $29,156	$0.67 - 2.7 billion
VAP	52,543	$19,633 - $28,508	$1 - 1.5 billion
CAUTI	449,334	$862 - $1,007	$1.2 - 1.6 billion

+SSI—surgical site infections; CLABSI—central venous line-associated bloodstream infection; VAP—ventilator-associated pneumonia; CAUTI—catheter-associated urinary tract infections.

Source: Scott RD. The direct medical costs of healthcare-associated infections in U.S. hospitals and benefits of prevention. Centers for Disease Control and Prevention, March 2009.

---

[d] Population groups often at greater risk for CA-MRSA include soldiers, athletic teams, prisoners, farmers, homeless, homosexual communities, population-dense households, nursing homes, and those in long-term rehabilitation.

Rapid Tests for MRSA

Microbiology laboratory tests provide clinicians with critical information to guide antimicrobial drug selection; implement control measures to prevent, isolate, or limit the spread of the infection; and monitor the effectiveness of treatment and control efforts.[15] Detecting MRSA-colonized patients is important for overall prevention of MRSA infections as well as control of MRSA transmission in health care settings. Newly colonized individuals have an estimated 30% risk of infection in the year following hospital discharge.[16]

Tests used for screening asymptomatic individuals typically use specimens from nasal swabs, whereas tests used for diagnosis use a blood or wound specimen. MRSA can be detected using newer culture methods such as broth-based enrichment, solid agar (salt-based) media, and chromogenic agar media, or molecular methods, including real-time polymerase chain reaction (PCR). The literature varies widely in the definition of rapid testing techniques—some researchers limit the term rapid testing to PCR-based methods, while others use a broader definition to include the newer culture-based methods. For purposes of this case study, the term "rapid tests" refers to those that detect MRSA in 24 hours or less; this applies to the newer culture-based methods and PCR-based methods. This does not include additional testing (or time) to identify the MRSA strain via culture-based methods.

Culture-based methods are highly accurate but may take multiple steps and more time to process. In general, there are two parts to culture-based MRSA testing: (1) confirmation of the presence or absence of MRSA; and (2) if the results are positive, susceptibility testing to identify the strain of MRSA present in the specimen. For example, for one type of culture test, specimens are obtained from patients[e] and set in a nutrient-rich liquid that facilitates growth of MRSA; the enrichment broth is sub-cultured to a solid selective medium that identifies the presence of MRSA. This first step takes at least 24 hours (laboratory time). The second step, antibiotic susceptibility testing, determines the antibiotics to which the micro-organisms are resistant and susceptible, and requires another 24 hours.[17] Total test turnaround time, counting laboratory, transport, and other time-related factors, can extend 48-72 hours.

More recently, innovative **PCR-based molecular techniques** have reduced detection time for certain pathogens, including MRSA, to a few (e.g., two) hours of laboratory time. However, susceptibility testing is still required to identify resistance and susceptibility to antibiotics using the techniques described above.[18] Molecular methods test for certain genetic components[f] of MRSA and consolidate certain aspects of testing. Some rapid MRSA tests are designed as diagnostic tests to identify MRSA only from sterile specimens (e.g., blood) whereas other tests were developed as screening tests that can detect MRSA from nonsterile nasal and other swabs.[19, 20] The high negative predictive value of these assays and their shortened turnaround time provide a valuable tool for infection control. Several studies have found increased use of PCR-based testing as the standard of care when compared to direct routine culture.[21-23]

Clinical Value of MRSA Testing

---

[e] Specimens obtained from individuals may be either nonsterile, e.g., swabs from such anatomical sites as nostrils or throat, or sterile, e.g., blood, cerebrospinal fluid, other body fluids.

[f] Most commercially available molecular methods detect the junction between the *S. aureus* chromosome and the SccMec cassette that carries the mecA gene.

The clinical analytic framework for MRSA testing depicts high clinical and economic value, as shown in Figure 1 (next page). Rapid MRSA testing yields **clinical value** in several different ways, including:

- **Enhanced capacity for active surveillance** to detect MRSA in asymptomatic individuals and minimize patient-to-patient transmission

- **Improved diagnostic accuracy and predictive value** in tests to identify individuals with true positive colonizations and infections and rule out true negatives

- **Improved therapeutic decision-making** for antibiotic selection and implementation of control precautions

- **Avoidance of unnecessary patient isolation** and the adverse psychosocial effects of isolation (i.e., depression and anxiety)

- **Potential for improved patient outcomes,** e.g., decreased mortality and morbidity and increased quality of life (e.g., less discomfort, faster recovery)

- **Decreased costs of care** as a result of decreased length of hospital stay and cost of treatments and management

Figure 1. Clinical Analytic Framework for Rapid MRSA Testing

Improved Health Outcomes

Universal surveillance with rapid MRSA tests can:
- Eliminate infection or colonization when used with an infection control program
- Decrease transmission rates by 75% in ICUs,
- 40% in non-ICU areas, and 67% hospital-wide[1]
- Identify newly acquired MRSA[2]

Rapid MRSA testing can:
- Reduce inappropriate use of antibiotics[3]
- Reduce likelihood of persistent carriage due to antibiotic failure

Decreased prevalence of MRSA from rapid testing, detection, and treatment can:
- Reduce incidence of depression and anxiety associated with isolation
- Decrease likelihood of adverse patient safety events associated with treatment (e.g., pressure ulcers, falls, electrolyte imbalances, fewer examinations of isolated patients)[4]

Improved Quality of Life

Decreased prevalence of MRSA from rapid testing, detection, and treatment can
- Quicken recovery, decreasing hospital and ICU length of stay
- Prevent unnecessary isolation due to unknown status[5-7]

Improved Cost Savings and Effectiveness
- Culture tests yield savings of $793 per hospital patient day; over 5 yrs, estimated costs are $34 million with 5,900 infections prevented and 42,000 patient days saved[8]

- PCR tests yield savings of $1,175 per hospital patient day; over 5 yrs, estimated costs are $40 million with 5,068 infections prevented and 33,90 patient days saved[8]

Sources:

1. Huang SS, Yokoe DS, Hinrichsen VL, et al. Impact of routine intensive care unit surveillance cultures and resultant barrier precautions on hospital-wide methicillin-resistant *Staphylococcus aureus* bacteremia. Clin Infect Dis 2006;43:971-8.
2. Harbarth S, Fankhauser C, Schrenzel J, et al. Universal screening for methicillin-resistant *Staphylococcus aureus* at hospital admission and nosocomial infection in surgical patients. JAMA 2008;299(10):1149-57.
3. Shales DM, Gerding DN, John JF, et al. Society for Healthcare Epidemiology of America and Infectious Diseases Society of America Joint Committee on the prevention of antimicrobial resistance: guidelines for the prevention of antimicrobial resistance in hospitals. Infect Control Hosp Epidemiol 1997;18:275-91.
4. Stelfox HT, Bates DW, Redelmeier DA. Safety of patients isolated for infection control. JAMA 2003;290(14):1899-905.
5. Shorr AF. Epidemiology and economic impact of methicillin-resistant *Staphylococcus aureus*: review and analysis of the literature. Pharmacoeconomics 2007;25:751-68.
6. Engemann JJ, Carmeli Y, Cosgrove SE, et al. Adverse clinical and economic outcomes attributable to methicillin resistance among patients with *Staphylococcus aureus* surgical site infection. Clin Infect Dis 2003;36:592-8.
7. Gould IM. Costs of hospital-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) and its control. Int J Antimicrob Agents 2006;28:379-84.
8. Consultation report on health technology. Clinical and cost effectiveness of screening for MRSA. Scotland: National Health Service, 2006. http://www.nhshealthquality.org/nhsqis/files/Consultation%20Final%20to%20Print.pdf.

Detection

Rapid MRSA tests have strong analytic validity.[g] A recent literature review by the Clinical and Laboratory Standards Institute reported that, for culture-based techniques, 95-96% of positive test results are correctly identified (*sensitivity*) and 96.6-98% of negative test results are correctly identified (*specificity*).[17] Culture methods also have high negative predictive value (NPV) at 99.2-100% and low false-positive rate at 1.8%; positive predictive value (PPV) is 76.8%.[17] For molecular-based methods, reported specificity of 95-96.4% and NPV of 98.2% are comparable to culture-based methods; sensitivity is 86.3-93.5% and PPV is 85.4%.[17, 24] Although the false-positive rate is slightly more (3.8%), the high sensitivity and NPV ensure very low risk of false-negative results.

This analytic validity confers significant value in detecting prevalence and new strains of MRSA in asymptomatic individuals via active surveillance. For example, one study conducted at an academic medical center during 1998-2002 sought to determine the proportion of patients who were colonized with MRSA at hospital admission (active surveillance) that were actually identified by microbiology cultures during hospitalization (routine[h] surveillance). The investigators found that testing only those patients with suspected infection underestimated hospital-wide prevalence of MRSA by 85%.[25] Rapid MRSA testing for active surveillance can also discern those individuals that entered the facility with MRSA from those that acquired MRSA during clinical care. Accordingly, CDC's clinical guidelines include recommendations for active surveillance[26] as do the recommendations recently released by a consortium[i] of health system stakeholders.[10]

Treatment

With rapid MRSA and susceptibility testing, antibiotic treatment can be targeted to the microorganism's susceptibility based on drug selection, dosing, and duration of treatment. This can prevent or minimize overuse of inappropriate antibiotics that can contribute to antibiotic resistance.[27] Targeted treatments enhance effective and efficient antibiotic management, allowing providers to monitor any changes in the relationship of resistance to specific drug-organism pairs and implement appropriate controls in antibiotic selection.

Colonized patients can begin immediate decolonization with topical antibiotic or antiseptic and systemic antibiotics to eliminate the pathogen or suppress transmission. One study reported a 52% reduction in incidence of MRSA cases among adult ICU patients after introducing an active surveillance program and decolonization regimen.[28] Contact precautions, including placing patients in private rooms and wearing of gloves, gown, and mask on entry into patients' rooms, can be applied immediately to reduce patient-to-patient transmission. Monitoring and

---

[g] Diagnostic accuracy is based on sensitivity and specificity. *Sensitivity* refers to how well a test identifies people with disease only among those who truly have the disease—most useful for ruling out disease when the test result is negative. *Specificity* refers to how well a test identifies people without disease only among those who truly do not have the disease—most useful for ruling in disease when a person tests positive. Positive and negative predictive value (PPV, NPV) estimate the probability of disease based on the prevalence of disease in the population.

[h] Routine (passive) surveillance refers to surveillance only of the specimens sent to the laboratory for testing.

[i] The consortium consists of: Society for Healthcare Epidemiology of America (SHEA), Infectious Diseases Society of America (IDSA), American Hospital Association (AHA), Association for Professionals in Infection Control and Epidemiology, Inc. (APIC), and The Joint Commission. Twenty-nine patient advocacy and professional organizations have endorsed the recommendations.

surveillance via patient retesting at regular intervals provide objective, scientific data for clearance to discontinue contact precautions.[26, 29]

Outcomes

Evidence is particularly strong for universal screening of high-risk or high-prevalence populations, such as those admitted to ICUs and those with previous infections.[30, 31] A nine-year study at Brigham and Women's Hospital (Boston), reported in 2006, found that routine surveillance cultures and subsequent contact precautions decreased incidence of bacteremia by 75% in ICUs, 40% in non-ICU areas, and 67% hospital-wide.[32] After issuing a mandate for MRSA testing in all high-risk units, the VA reported in 2007 that its Palo Alto site had reduced infection rates in ICUs by 79%.[33] A recently published study also provides evidence of the impact of universal screening of all hospital patients upon admission. A two-year study (2005-2007) of expanded surveillance of hospital patients found that MRSA prevalence decreased 70% with universal screening and 36% with ICU-only patient screening.[34]

In contrast, another two-year study (2004-2006) involving screening of all surgical patients (more than 10,000) admitted to a Swiss teaching hospital for more than 24 hours did not reduce MRSA infection rates (compared to the control group (also more than 10,000 surgical patients), although 337 previously unknown MRSA carriers were identified.[35] Another study of universal MRSA screening conducted in a London teaching hospital (1,200 beds) compared the effect of rapid PCR-based testing and culture-based testing on MRSA acquisition.[36] Study results indicated that rapid MRSA testing reduced reporting time and unnecessary, pre-emptive isolation days, but did not significantly reduce MRSA acquisition (3.2% vs. 2.8% for control group). While the impact of screening on MRSA reduction rates often can be difficult to quantify, the value to detection in asymptomatic individuals and quality of care are substantial. Larger trials are necessary to accurately determine the true magnitude of the reduction.

To date, few studies have examined the direct effect of rapid MRSA testing (targeted or universal) on mortality via shorter turnaround in obtaining culture reports and more prompt treatment. A study reported in 2000 conducted at the VA Maryland Health System found that up to 55% of patients received ineffective antibiotics in the first 48 hours while waiting for test results, indicative of wasted resources and poor quality and safety of care, although there was no effect on mortality.[37] Another study reported in 2003 that was conducted at a trauma center found that, for patients testing positive, a delay in treatment was an independent predictor of mortality in 33% of patients compared to 19% of patients who received earlier treatment (no delay).[38] The nine-year study at a teaching hospital in Boston (mentioned above) assessed MRSA prevalence among the 43,000 patients admitted annually (using culture-based methods) and found no difference in mortality rates between patients with MRSA and those with MSSA.[32] However, differences in study definitions, methodology, and small sample sizes contributed to variation in results.

Economic Value of Rapid MRSA Testing

Culture-based MRSA testing has demonstrated significant economic value, however, less is known about the economic value of rapid MRSA testing. Because rapid MRSA testing facilitates earlier and more accurate detection, it has the potential to yield economic value, in terms of:

> - Decreased treatment expenditures per case
> - Reduced overall hospital and health system costs
> - Improved cost-effectiveness per patient hospital day

Quantifying the health and economic impact of MRSA is essential to plan and justify resource allocation to counteract antibiotic resistance and prevent the spread of resistant organisms in the health care environment.[14] However, currently, there are few studies that document direct cost savings from use of rapid MRSA tests. Often, savings are inferred based on costs associated with HAIs more broadly and the decreased prevalence of MRSA and associated treatment costs.

A 2007 study at the Evanston Northwestern Healthcare (a three-hospital organization near Chicago) assessed the impact of a MRSA infection reduction program, including financial impact. The net cost (i.e., in addition to existing costs) of the universal MRSA screening program was $600,000 per year, with principle costs attributed to laboratory supplies and personnel. Contact isolation in those testing positive was considered part of the usual costs of the infection control practice and not the part of the screening and surveillance program. The overall number of patients testing positive for MRSA increased 20%, which translated to an annual expense of $44,000 or $15-16 per admission. Additional calculations of cost reductions were based on data from patients who remained in the hospital for at least 8 days, the median time to patient development of a MRSA infection. In this comparison, patients with MRSA infection had an excess medical expense of $23,783. Eliminating 50 reported infections during the first year resulted in a reduction of nearly $1.2 million in medical expenditures; after subtracting the $600K investment, this leaves the expanded screening program cost neutral. Along with this positive financial impact, the expanded screening program resulted in an 80% reduction in incidence of MRSA bloodstream infections at the end of the first year of the program.[39]

In 2006, the U.K. National Health Service conducted an extensive cost-effectiveness analysis of different MRSA testing methods (e.g., swab-based broth enrichment culture, real-time PCR) in hospital patients based on an estimated colonization rate of 7%.[40] Converting to U.S. inflation-adjusted dollars (2008), use of the culture-based method was associated with the lowest cost per patient (bed) day saved at $793, rendering it the more cost-effective method. Over five years, the culture-based method was estimated to cost a total of $34 million, prevent over 5,900 infections, save more than 42,000 bed days, and reduce MRSA prevalence to 1.4% of all HAIs. The PCR-based method had a higher cost per bed day saved at $1,175. Over five years, the PCR strategy was estimated to cost $40 million, prevent 5,068 infections, save 33,900 bed days, and result in a MRSA prevalence rate of 1.7% of all HAIs, indicating slightly less cost effectiveness than conventional cultures. For either method, swabbing only high-risk patients had a comparable cost per bed day saved of $793, but only reduced MRSA prevalence to 5.3%.

However, the economic model described above was not predictive and the outcome values should not be interpreted as absolute. In addition, there were several limitations in the parameters used in the modeling study—namely, uncertainty about the duration of MRSA carriage within the community; uncertainty about the transmission behavior of hospital-acquired MRSA strains within the community; uncertainty regarding hospital re-admission patterns in the population; uncertainty concerning the effectiveness of MRSA decolonization treatment; the unknown prevalence of MRSA colonization within the population; and the impact of other infectious diseases competing for isolation facilities. Furthermore, the simplicity of the model may not fully capture the benefits of PCR-based testing.

The findings of a study presented in 2008 were based on a computerized simulation that assessed the clinical and economic impact of five surveillance strategies for MRSA control in ICUs.[41] An individual-based mathematical model was developed to simulate transmission of MRSA in a 10-bed ICU over a one-year period in one-hour increments. The five surveillance strategies simulated included: (1) passive surveillance (isolation of only those patients with a history of MRSA colonization or infection), (2) active surveillance using anterior nares cultures, with isolation when cultures return positive in a mean of 48 hours, (3) active surveillance using a PCR-based test with a 24-hour return time, (4) active surveillance using a PCR-based test with an 8-hour return time and (5) active surveillance using a rapid PCR-based test with a 2-hour return time. All simulations assumed patients would be isolated only when tests returned positive. Base-case analyses[j] assumed median length of stay of 2 days (mean 4.6 days), 12.3% MRSA colonized on admission, and 90% compliance with obtaining cultures.[41] Costs of tests[k] and isolation/control measures (gown/glove/time) were included in calculations.

Table 2 summarizes study results/predictions based on 100 model runs with 776 average yearly admissions. From the hospital perspective, all active surveillance strategies were cost-effective. As the speed of detection improves from 48 hours to 2 hours, the number of cases prevented increases. MRSA acquisition rates were lowest with the 2-hr PCR-based test at 5.9% compared to the other strategies at 6.2-9.0%. PCR-based tests also showed a significant improvement in test turnaround time, providing test results prior to discharge for 99% of patients with the 2-hr PCR test and 95% of patients with the 8-hr PCR test compared to 71% with the 24-hr PCR test and only 50% with the culture-based test. While cost effectiveness is evident in all but the passive surveillance strategy, the culture-based test and 24-hr PCR test demonstrated the highest potential savings per MRSA acquisition and bacteremia prevented. This difference is most likely attributable to the lower cost per test of cultures and the 24-hr PCR test as compared to the 8-hr and 2-hr PCR tests. The investigators concluded that, if the attributable hospital cost burden of a preventable MRSA bacteremia is greater than $17,000, then each strategy would be cost-saving from the hospital perspective.[41]

---

[j] Parameter estimates were derived from the literature and an active surveillance database from a 3-year cohort of patients admitted to a medical ICU of a tertiary-care hospital.

[k] Using 2006 data, standard cultures cost $11.73 if positive and $7.73 if negative, while PCR-based tests cost $37.01 for 8-hour and 24-hour return and $43.66 for rapid-PCR with 2-hour return.

Table 2. Clinical and economic benefit of five surveillance strategies for MRSA control in ICUs

	Passive surveillance	Active surveillance - Culture (48 hr return)	Active surveillance PCR (24-hr return)	Active surveillance PCR (8-hr return)	Active surveillance PCR (2-hr return)
Clinical benefit					
MRSA acquisition rate	9%	7.2%	6.7%	6.2%	5.9%
Total annual patient days in isolation (all patients)	230	449	510	529	552
Test results returned prior to patient discharge from ICU	n/a	50%	71%	92%	99%
Cost-effectiveness					
Costs per new MRSA colonization prevented	n/a	$738	$1,612	$1,391	$1,300
Costs per MRSA bacteremia prevented	n/a	$10,083	$16,968	$14,603	$14,825

In general, it is difficult to draw overarching conclusions about cost effectiveness and cost savings of conventional and rapid MRSA testing because of the current paucity of high-quality economic evaluations designed to produce direct evidence.[40] In real-world clinical practice, there is a trade-off between the increased timeliness of PCR-based MRSA testing which supports faster implementation of infection control procedures and the higher cost of the test and slightly lower test sensitivity compared to culture-based methods. In the near term, it is unlikely that PCR assays will fully replace culture-based methods, especially for blanket screening; yet, certain studies do support a role for rapid PCR-based MRSA testing on specific hospital wards where risk of colonization or infection is high and preadmission testing is not possible. Current clinical evidence suggests the potential for significant cost savings with rapid MRSA testing due to decreases in hospital length of stay, morbidity, and other costs.

Conclusions

The health and economic burden of HAIs is substantial and new strains resistant to antibiotics are evolving rapidly. Advances in testing techniques have led to substantial improvements in test sensitivity and specificity as well as decreases in detection time to 24 hours for culture-based methods and a few hours for PCR-based methods.

Public and private sector stakeholders recognize the significance of the public health threat of HAIs, including MRSA and other highly resistant strains, and are increasingly requiring health care institutions to implement proactive screening strategies that can reduce prevalence and transmission. Screening strategies examined in the literature include both targeted screening of specific patient subpopulations and universal screening of all individuals upon hospital admission. Evidence as to the best strategy is mixed. Evidence is particularly strong supporting screening of high-risk or high-prevalence populations, such as those admitted to ICUs and those with previous infections. When expanded to universal screening of all

individuals admitted to the hospital, evidence varies. Some studies report 70-80% decreases in prevalence of MRSA whereas other studies report little or no change in prevalence of MRSA. Nevertheless, all studies confirmed that there was substantial value in identifying undetected carriers of MRSA to provide appropriate treatment and potential for transmission. Additional large-scale studies are needed to determine statistically significant effects of universal screening on prevalence and mortality.

The few studies on the economic impact of MRSA testing that have been conducted indicate the potential for significant savings, particularly when considering direct medical costs of up to $45 billion per year for all HAIs, including for costs associated with increased hospital/ICU length of stay, use of single rooms, and use of more expensive antibiotics. Although the investment in implementing an expanded screening program can be somewhat high, the potential savings and return on investment may be substantial given the expected decreases in direct medical costs. However, studies that specifically calculate reductions in direct medical costs as a result of rapid MRSA testing are needed to support these expectations. Because of the importance of HAIs and MRSA to patient safety, quality of care, and public health, it is essential that stakeholders ensure that the appropriate policies are in place to support broad implementation of active screening and surveillance programs to detect antibiotic resistance in a growing proportion of the population.[9, 10]

References

1. Antimicrobial resistance: issues and options. Harrison PF, Lederberg J, eds. Washington, DC: National Academy Press, 1998.

2. Patient safety: health care-associated infections. Agency for Healthcare Research and Quality, 2008. (Accessed April 15, 2009, at http://www.ahrq.gov/qual/haiflyer.pdf.)

3. Scott RD. Direct medical costs of healthcare-associated infections in U.S. hospitals and the benefits of prevention. Atlanta: Centers for Disease Control and Prevention, 2009.

4. Raygada JL, Levine DP. Methicillin-resistant *Staphylococcus aureus*: a growing risk in the hospital and in the community. American Health and Drug Benefits 2009;2(2):86-95.

5. Boyce JM, Jackson MM, Pugliese G, et al. Methicillin-resistant *Staphylococcus aureus* (MRSA): a briefing for acute care hospitals and nursing facilities. The AHA Technical Panel on Infections Within Hospitals. Infection Control and Hospital Epidemiology 1994;15:105-15.

6. Hidron AI, Edwards JR, Patel J, et al. (2008) NHSN annual update. Antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006-2007. Infection Control and Hospital Epidemiology 2008;29(11):997-1011.

7. Nguyen HH, Hoze MD. Hospital-based outpatient parenteral antimicrobial therapy (OPAT) at a university hospital. Presented at the 44th Annual Meeting of the Infectious Diseases Society of America, October 12-15, 2006, Toronto, Canada.

8. Jones JB, Stewart WF. Methicillin-resistant *Staphylococcus aureus*: epidemiology and antimicrobial use. American Health and Drug Benefits 2009;2(2 (Supplement)):S68-S73.

9. Building a surveillance framework for drug-resistant *Streptococcus pneumoniae* and methicillin-resistant *Staphylococcus aureus*. Atlanta: Centers for Disease Control and Prevention, 2003.

10. Calfee DP, Salgado CD, Classen D, et al. Strategies to prevent transmission of methicillin-resistant *Staphylococcus aureus* in acute care hospitals. Infection Control and Hospital Epidemiology 2008;29(Supplement 1):S62-S80.

11. Shorr AF. Epidemiology and economic impact of methicillin-resistant *Staphylococcus aureus*: review and analysis of the literature. Pharmacoeconomics 2007;25:751-68.

12. Engemann JJ, Carmeli Y, Cosgrove SE, et al. Adverse clinical and economic outcomes attributable to methicillin resistance among patients with *Staphylococcus aureus* surgical site infection. Clin Infect Dis 2003;36:592-8.

13. Gould IM. Costs of hospital-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) and its control. Int J Antimicrob Agents 2006;28:379-84.

14. Cosgrove SE, Qi Y, Kaye KS, Harbarth S, Karchmer AW, Carneli Y. The impact of methicillin resistance in *Staphylococcus aureus* bacteremia on patient outcomes: mortality, length of stay, and hospital charges. Infection Control and Hospital Epidemiology 2005;26:166-74.

15. Lichtenberg FR. The value of diagnostics: the impact of hospital antimicrobial susceptibility testing. Washington, DC: AdvaMed, 2008.

16. Jernigan JA. Prevention and control of healthcare-associated methicillin-resistant *Staphylococcus aureus*. 2008. Atlanta, Centers for Disease Control and Prevention.

17. Screening for methicillin-resistant *Staphylococcus aureus*: principles, practices, and potential problems; a report. Clinical and Laboratory Standards Institute, 2008.

18. Cai F, Chen H, Hicks CB, Bartlett JA, Zhu J, Gao F. Detection of minor drug-resistant populations by parallel allele-specific sequencing. Nature Methods 2007;4:123-5.

19. Levi K, Towner KJ. Detection of methicillin-resistant *Staphylococcus aureus* (MRSA) in blood with the EVIGENE MRSA detection kit. J Clin Microbiol 2003;41(8):3890-2.

20. Huletsky A, Lebel P, Picard FJ, et al. Identification of methicillin-resistant *Staphylococcus aureus* carriage in less than 1 hour during a hospital surveillance program. Clinical Infectious Disease 2005;40:976-81.

21. Alonzo TA, Pepe MS. Using a combination of reference tests to access the accuracy of a new diagnostic test. Stat Med 1999;18:2987-3003.

22. Boyce JM, Havill NL. Comparison of BD GeneOhm methicillin-resistant *Staphylococcus aureus* (MRSA) PCR versus the CHROMagar MRSA assay for screening patients for the presence of MRSA strains. J Clin Microbiol 2008;46:350-1.

23. Carroll KC. Rapid diagnostics for methicillin-resistant *Staphylococcus aureus*: current status. Molecular Diagnostics and Therapeutics 2008;12:15-24.

24. wolk DM, Picton E, Johnson D, et al. A multicenter evaluation of the Cepheid Xpert MRSA test as a rapid screening method for detection of methicillin-resistant *Staphylococcus aureus* (MRSA) from nares swabs. J Clin Microbiol;47:758-64.

25. Salgado CD, Farr BM. What proportion of hospital patients colonized with methicillin-resistant *Staphylococcus aureus* are identified by clinical microbiology cultures? Infection Control and Hospital Epidemiology 2006;27:116-21.

26. Siegel JD, Rhinehart E, Jackson M, Chiarello L. Management of multidrug-resistant organisms in healthcare settings, 2006. Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, 2006.

27. Shales DM, Gerding DN, John JF, et al. Society for Healthcare Epidemiology of America and Infectious Diseases Society of America Joint Committee on the Prevention of Antimicrobial Resistance: guidelines for the prevention of antimicrobial resistance in hospitals. Infection Control and Hospital Epidemiology 1997;18:275-91.

28. Ridenour C, Lampen R, Federspiel J, et al. Selective use of intranasal mupirocin and chlorhexidine bathing and the incidence of methicillin-resistant *Staphylococcus aureus* colonization and infection among intensive care unit patients. Infection Control and Hospital Epidemiology 2007;28:1155-61.

29. Wilson SJ, Knipe CJ, Zieger MJ, et al. Direct costs of multidrug-resistant *Acinetobacter baumannii* in the burn unit of a public teaching hospital. Am J Infect Control 2004;32(6):342-4.

30. Huang SS, Platt R. Risk of methicillin-resistant *Staphylococcus aureus* infection after previous infection or colonization. Clin Infect Dis 2003;43:971-8.

31. West T, Guerry C, Hiott M, et al. Effect of targeted surveillance for control of methicillin-resistant *Staphylococcus aureus* in a community hospital system. Infection Control and Hospital Epidemiology 2006;27:233-8.

32. Huang SS, Yokoe DS, Hinrichsen VL, et al. Impact of routine intensive care unit surveillance cultures and resultant barrier precautions on hospital-wide methicillin-resistant *Staphylococcus aureus* bacteremia. Clin Infect Dis 2006;43:971-8.

33. VHA Directive 2007-002. Methicillin-resistant *Staphylococcus aureus* (MRSA) initiative. Veterans Health Administration, 2007.

34. Robicsek A, Beaumont JL, Paule SM, et al. Universal surveillance for methicillin-resistant *Staphylococcus aureus* in 3 affiliated hospitals. Ann Intern Med 2008;148:409-18.

35. Harbarth S, Fankhauser C, Schrenzel J, et al. Universal screening for methicillin-resistant *Staphylococcus aureus* at hospital admission and nosocomial infection in surgical patients. JAMA 2008;299(10):1149-57.

36. Jeyaratnam D, Whitty C, Phillips K, et al. Impact of rapid screening tests on acquisition of methicillin resistant *Staphylococcus aureus*: cluster randomized crossover trial. Br Med J 2008;336(7650):927-30.

37. Roghmann MC. Predicting methicillin resistance and the effect of inadequate empiric therapy on survival in inpatients with *Staphylococcus aureus* bacteremia. Arch Intern Med 2000;160:1001-4.

38. Lodise TP, McKinnon PS, Swiderski L, Rybak MJ. Outcomes analysis of delayed antibiotic treatment for hospital-acquired *Staphylococcus aureus* bacteremia. Clin Infect Dis 2003;36:1418-23.

39. Peterson LR, Hacek DM, Robicsek A. 5 Million Lives Campaign. Case study: an MRSA intervention at Evanston Northwestern Healthcare. The Joint Commission Journal on Quality and Patient Safety 2007;33(12):732-8.

40. Consultation report on health technology. Clinical and cost effectiveness of screening for MRSA. Scotland: National Health Service, 2006. http://www.nhshealthquality.org/nhsqis/files/Consultation%20Final%20to%20Print.pdf.

41. Perencevich EN, Furuno JP, Bradham DD, Hartley DM, Harris AD, Gaff HD. Clinical and Economic Impact of Five Surveillance Strategies, Including Rapid-PCR Testing, for Methicillin-resistant Staphylococcus aureus (MRSA) Control in Intensive Care Units (ICUs). Society for Healthcare Epidemiology of America, 2008. (Accessed September 20, 2009, at http://www.shea-online.org/Assets/files/Sunday_Afternoon_Submitted_Scientific_Papers.pdf.)

Case Study II:

Screening for Type 2 Diabetes and Prediabetes with HbA1c

Magnitude and Importance of Diabetes and Prediabetes

Diabetes includes a group of chronic diseases marked by abnormally high levels of blood glucose resulting from defects in insulin production, insulin action, or both. The American Diabetes Association (ADA) estimates that close to 24 million American children and adults, 7.8% of the population, have diabetes.[1] Of these, about 17.9 million are diagnosed and 5.7 million are undiagnosed. Ninety percent of all diabetes is type 2,[a] an acquired form of the disease with adult onset.[1, 2] In 2007, 1.6 million new cases of diabetes were diagnosed in adults age 20 years and older.[1]

The complications of diabetes have a substantial impact on mortality and morbidity, as well as costs.[2] According to the CDC, diabetics have a two-to four-fold increase in heart disease-related mortality and risk of stroke relative to nondiabetics. Diabetes is the leading cause of new cases of blindness, kidney failure, and non-traumatic lower-limb amputations. About 60-70% of diabetics have mild-to-severe nervous system damage and 70% have hypertension. Other complications include dental disease, risks to pregnancy, acute life-threatening biochemical imbalances and increased susceptibility to other illnesses.[2]

The total annual cost of diabetes in the U.S. was an estimated $174 billion in 2007, including $116 billion in direct health care costs, $58 billion of which was for chronic diabetes-related complications. About $58 billion of total costs of diabetes were attributable to indirect costs (lost productivity), including increased absenteeism, reduced performance at work, unemployment disability, and premature death.[3] Diabetes is among the most costly health conditions for employers in terms of direct and indirect costs.[4]

In addition to those diagnosed with diabetes, an estimated 57 million individuals are prediabetic for type 2.[1, 2] Prediabetes is an asymptomatic state characterized by insulin resistance, impaired glucose tolerance (IGT) or fasting glucose (IFG), or other endocrine and metabolic conditions. Individuals with prediabetes have glucose levels higher than normal but slightly lower than those classified as diabetic. Because type 2 diabetes typically progresses slowly over an extended time of 10-12 years, prediabetic individuals are at increased risk of developing microvascular and macrovascular complications prior to diagnosis.[5] Often, these complications already are present at the time of clinical diagnosis when signs and symptoms of diabetes are present.

The single greatest contributing factor to the onset of type 2 diabetes is obesity attributed to poor diet and sedentary lifestyle, which can be corrected with appropriate screening, patient education, and behavioral interventions.[1, 2] Other risk factors include older age (≥65 years),

---

[a] Type 2 diabetes usually begins as insulin resistance, a disorder in which body cells do not respond appropriately when insulin is present. As the need for insulin arises, the pancreas gradually loses its ability to produce it. Type 2 is managed with a combination of dietary treatment, exercise, medications and, if necessary, insulin supplementation. Type 1 diabetes is a genetic-based, insulin-dependent disorder affecting 5-10% of all those diagnosed.

family history, and race/ethnicity (e.g., higher prevalence among African Americans, Latino Americans, Native Americans, and Asian-Pacific Islanders).[1,2] Family and primary care physicians have opportunities to identify those at high risk for developing diabetes and instituting primary prevention strategies. Analyses documenting increased use of primary care services and prescriptions for various health conditions in the five-year period leading up to diagnosis of diabetes suggest that there are numerous opportunities for diabetes screening and earlier intervention.[6]

Given the individual and population impact of diabetes and prediabetes, there is substantial opportunity to identify these conditions early and accurately prediabetic individuals and initiate appropriate preventive and therapeutic interventions. In recent years, certain well-designed randomized-controlled clinical trials (RCTs) and a growing body of observational evidence have demonstrated the benefits of early lifestyle and/or therapeutic interventions in delaying or preventing the onset of diabetes and related complications.[7-10] However, no RCTs have been done to establish whether routine screening for diabetes and prediabetes in asymptomatic individuals improves health outcomes compared to clinical diagnosis of symptomatic individuals.[11,12] Several observational studies (discussed below) suggest clinical value of routine screening of asymptomatic individuals.

In its 2008 update, the U.S. Preventive Services Task Force (USPSTF) concluded that there is still insufficient clinical trial data[b] to recommend for or against routine screening of asymptomatic individuals with blood pressure of 135/80 mm Hg or lower for type 2 diabetes, IGT, or IFG, effectively leaving decision making to clinicians.[11] USPSTF maintained its earlier recommendation for screening of type 2 diabetes in asymptomatic individuals with sustained blood pressure (either treated or untreated) greater than 135/80 mm Hg.

Laboratory Tests for Diabetes and Prediabetes

Conventional laboratory tests that are used to identify diabetic or prediabetic individuals measure the impaired glucose tolerance (IGT) or impaired fasting glucose (IFG) that are characteristic of these conditions. Cutoff points in test results that have been determined by the ADA for the diagnoses of diabetes and prediabetes are the following:

> - Prediabetes: IGT (2-hour postprandial [after a meal] glucose) of 140-199mg/ml or IFG (fasting glucose) of 100-125mg/ml, or both.
>
> - Diabetes: IGT (2-hour postprandial glucose) of ≥ 200mg/ml or IFG (fasting glucose) ≥ 126mg/ml, or both.

Source: Zhang P, Engelgau MM, Valdez R, et al. Efficient cutoff points for three screening tests for detecting undiagnosed diabetes and prediabetes. Diabetes Care 2005;28(6):1321-5.

---

[b] In 2008, the Evidence-based Practice Center of the Oregon Health & Science University and Portland Veterans Administration Medical Center (Oregon EPC) completed a systematic review of evidence on screening adults for type 2 diabetes and prediabetes in primary care settings, focusing on literature published since 2003.[12] This review was conducted in support of a 2008 update by the USPSTF of its 2003 recommendations on diabetes screening.[11]

The gold standard test for diagnosing diabetes is the **oral glucose tolerance test (OGTT)** (2-hour), which requires an 8-hour fast and administration of oral glucose to determine how quickly it is cleared from the blood. Blood is drawn at designated intervals, including the 2-hour draw, which is the most important for diagnosing diabetes. Although the OGTT is more accurate than other tests for diagnosis, it also is more cumbersome, time consuming, and costly.

The most widely used screening test for diabetes is the **fasting plasma glucose test (FPG).** Because FPG requires an 8-hour fast and only one blood draw, it is a faster test, more acceptable to patients, and less expensive than OGTT. FPG results are more reproducible than the OGTT but less sensitive in detecting true positives, so a second, confirmatory test (either FPG or OGTT) is needed to verify positive results. Providers may order a **random plasma glucose test (RPG)** given with minimal (few hours) or no fasting; however, only individuals with severely high glucose levels can be identified in this manner for further testing as diet, medications, and other factors influence RPG test results.

The **hemoglobin A1c test[c] (HbA1c)** reflects average glycemia (blood glucose levels) during the preceding 2-3 months with a single blood draw; no fast is required. HbA1c is the standard test for monitoring glycemic levels once diagnosis of diabetes is established with OGTT or FPG tests. There is significant interest in using HbA1c as a screening tool because it is easier to obtain than FPG, requiring just one blood sample without fasting or capillary blood glucose (CBG) sample for a point of care device.[d] HbA1c is considered to be an effective 'hybrid' of OGTT and FPG.[14] Because HbA1c and FPG do not involve a glucose challenge, both tests require a confirmation test. IGT can only be confirmed with OGTT.

Since 1993, clinical laboratories have worked to standardize HbA1c values among the different test systems. Recently, an international consortium was created to establish world-wide standardization and harmonization of HbA1c measurement, in order to facilitate use of HbA1c as a diagnostic tool for diabetes. In its 2008 update on recommendations pertaining to diabetes screening, the USPSTF noted that the utility of HbA1c testing was limited in the past by its relatively poor reproducibility and lack of standardization across laboratories, but that widespread adoption of standardized hemoglobin A1c measurements has occurred recently and that the current techniques are generally highly reproducible.[11]

In early 2010, ADA is expected to issue guidelines confirming use of HbA1c as a valid diagnostic test for diabetes. Currently, diagnostic cutoff points are being redefined from previous values: HbA1c of 4.0-6.4% as nondiabetic and 6.5% as diabetic. In June 2009, members of an International Expert Committee who were appointed by the ADA, the European Association for the Study of Diabetes, and the International Diabetes Federation recommended that an HbA1c of 6.5% be the threshold for diagnosing diabetes.[15, 16] The committee also recommended that patients with HbA1c levels between 6.0% and 6.5% be cautioned that they are in the highest risk group for developing type 2 diabetes.

---

[c] HbA1c tests include central laboratory tests and point of care test systems.

[d] Random capillary blood glucose (CBG) testing relies on a glucose meter that patients can use to self-monitor glucose levels or that providers can use at the point of care. Fasting is not required, and a superficial skin prick provides a drop of blood for the test. Because CBG test results can differ by as much as 10-15% from reference laboratory values, CBG testing is not considered valid for diagnosing diabetes.[13]
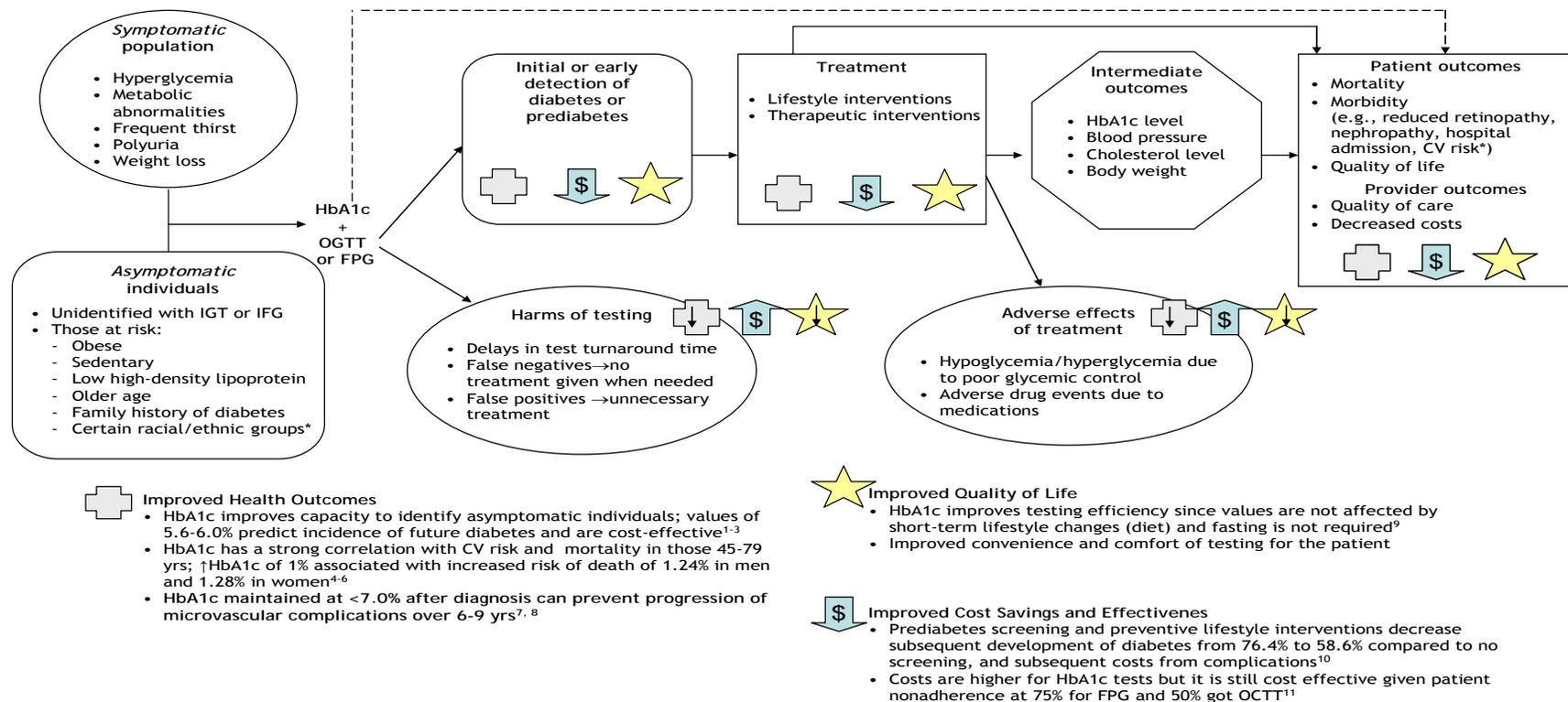
Clinical Value of Diabetes and Prediabetes Screening with HbA1c

The clinical and economic value of screening asymptomatic individuals for diabetes and prediabetes is summarized in the clinical analytic framework depicted in Figure 1. Use of HbA1c for screening of asymptomatic individuals and diagnosis of symptomatic individuals supports early detection and accurate diagnosis. In turn, this informs clinical decision-making regarding lifestyle and therapeutic interventions. Testing and treatment, including ongoing monitoring, can yield improvements in patient outcomes, including:

- **Improves capacity to identify** asymptomatic individuals with sustained elevated glucose levels
- **Aligns detection** of the development of cardiovascular disease and microvascular complications (retinopathy, nephropathy, neuropathy)
- **Has comparable diagnostic accuracy and predictive values** to certain other tests currently used
- **Improves testing efficiency** since values are not affected by short-term lifestyle changes (diet) and fasting is not required
- **Improves convenience and comfort** of testing for the patient
- **Improves clinical decision-making** for early lifestyle and therapeutic interventions
- **Improves patient outcomes**, including decreased mortality and morbidity and increased quality of life
- **Decreased costs of care** as a result of diabetes prevention, delay of onset, and/or reduction in and severity of complications

Evidence supporting the impact of screening on health outcomes and costs or cost effectiveness is largely indirect, i.e., not based on prospective clinical trials tracking populations from screening through outcomes, but rather on a chain of evidence from separate studies leading from screening to outcomes.

Figure 1. Clinical Analytic Framework for Screening and Diagnosis of Prediabetes & Diabetes using HbA1c

Symptomatic population
- Hyperglycemia
- Metabolic abnormalities
- Frequent thirst
- Polyuria
- Weight loss

Asymptomatic individuals
- Unidentified with IGT or IFG
- Those at risk:
  - Obese
  - Sedentary
  - Low high-density lipoprotein
  - Older age
  - Family history of diabetes
  - Certain racial/ethnic groups*

HbA1c + OGTT or FPG

Initial or early detection of diabetes or prediabetes

Treatment
- Lifestyle interventions
- Therapeutic interventions

Intermediate outcomes
- HbA1c level
- Blood pressure
- Cholesterol level
- Body weight

Patient outcomes
- Mortality
- Morbidity (e.g., reduced retinopathy, nephropathy, hospital admission, CV risk*)
- Quality of life

Provider outcomes
- Quality of care
- Decreased costs

Harms of testing
- Delays in test turnaround time
- False negatives→no treatment given when needed
- False positives →unnecessary treatment

Adverse effects of treatment
- Hypoglycemia/hyperglycemia due to poor glycemic control
- Adverse drug events due to medications

Improved Health Outcomes
- HbA1c improves capacity to identify asymptomatic individuals; values of 5.6-6.0% predict incidence of future diabetes and are cost-effective[1-3]
- HbA1c has a strong correlation with CV risk and mortality in those 45-79 yrs; ↑HbA1c of 1% associated with increased risk of death of 1.24% in men and 1.28% in women[4-6]
- HbA1c maintained at <7.0% after diagnosis can prevent progression of microvascular complications over 6-9 yrs[7, 8]

Improved Quality of Life
- HbA1c improves testing efficiency since values are not affected by short-term lifestyle changes (diet) and fasting is not required[9]
- Improved convenience and comfort of testing for the patient

Improved Cost Savings and Effectivenes
- Prediabetes screening and preventive lifestyle interventions decrease subsequent development of diabetes from 76.4% to 58.6% compared to no screening, and subsequent costs from complications[10]
- Costs are higher for HbA1c tests but it is still cost effective given patient nonadherence at 75% for FPG and 50% got OCTT[11]

*CV is cardiovascular risk

Sources:
1. Edelman D, Olsen MK, Dudley TK, et al. Utility of hemoglobin A1c in predicting diabetes risk. J Gen Intern Med 2004;19(12):1175-80.
2. Norberg M, Eriksson JW, Lindahl B, et al. A combination of HbA1c, fasting glucose and BMI is effective in screening for individuals at risk of future type 2 diabetes. J Intern Med 2006;260(3):263-71.
3. Norris SL, Kansagara D, Bougatsos C, Fu R. Screening adults for type 2 diabetes: a review of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2008;148:855-68.
4. Khaw KT, Wareham N, Bingham S, et al. Association of hemoglobin A1c with cardiovascular disease and mortality in adults: the European Prospective Investigation into Cancer in Norfolk. Ann Intern Med 2004;141:413-20.
5. Gerstein HC. Glycosylated hemoglobin: finally ready for prime time as a cardiovascular risk factor. Ann Intern Med 2004;141:475-6.
6. Selvin E, Marinopoulos S, Berkenblit G, et al. Meta-analysis: glycosylated hemoglobin and cardiovascular disease in diabetes mellitus. Ann Intern Med 2004;141:421-31.
7. Ohkubo Y, Kishikawa H, Araki E, et al. Intensive insulin therapy prevents the progression of diabetic microvascular complications in Japanese patients with non-insulin-dependent diabetes mellitus: a randomized prospective 6-year study. American College of Physicians Journal Club 1996;123(3):68.
8. Tanaka Y, Atsumi Y, Matsuoka K, et al. Role of glycemic control and blood pressure in the development and progression of nephropathy in elderly Japanese NIDDM patients. Diabetes Care 1998;21(1):116-20.
9. Tahara Y, Shima K. The response of GHb to stepwise plasma glucose change over time in diabetic patients. Diabetes Care 1993;16(9):1313-4.
10. Hoerger TJ, Hicks KA, Sorensen SW, et al. Cost-effectiveness of screening for prediabetes among overweight and obese U.S. adults. Diabetes Care 2007;30(11):2874-9.
11. Zhang P, Engelgau MM, Valdez R. Costs of screening for prediabetes among U.S. adults: a comparison of different screening strategies. Diabetes Care 2003;26(9):2536-42.

Detection

A recent systematic review compared the diagnostic accuracy of HbA1c and FPG in detecting diabetes and prediabetes as confirmed by reference standard OGTT results according to World Health Organization criteria.[17] According to the review, HbA1c and FPG are equally effective for detection of type 2 diabetes, with the recommended cutoff point for HbA1c at >6.1%. Certain studies also have demonstrated that HbA1c has less intra-individual variation and is a better predictor of micro- and macrovascular complications, whereas other studies indicate that HbA1c, OGTT, and FPG are equivalent predictors of the development of retinopathy and nephropathy.[18-20] Differences in study findings are partly attributed to the assessment of different cutoff points for each test. For example, HbA1c cutoff points can be set to optimize sensitivity or specificity to prediabetes or undiagnosed diabetes.

- For HbA1c, using a cutoff point of 5.9% to identify types 2 diabetes *sensitivity* (to determine true positives) is reported to be in the range of 76-95% and *specificity* (to determine true negatives) is 67-86%.[21, 22] When aligning optimum cutoff values with the Diabetes Control and Complications Trial (DCCT) at 6.1-6.2%, corresponding *sensitivity* is 43%-81% and *specificity* is 79%-99%.[20, 22-24]

- For FPG, using a cutoff at 5.6% to identify type 2 diabetes, sensitivity is reported to be in the range of 80-88% and specificity is 79.2-85.8%.[22, 25] A cutoff ≥ 6.1% is reported to yield sensitivity of 48-64% and specificity is 94-98%.[20, 24]

According to some good-quality studies cited in the evidence review by the Oregon EPC, HbA1c values in the range of 5.6-6.0% appear to predict a higher incidence of future diabetes in prediabetic patients and seem to be the most cost-effective for diagnosing type 2 diabetes.[12, 26, 27] However, the relationship between cut-off values and cost effectiveness can vary depending on whether the goal is to determine individual or population health status. For example, a cost-effectiveness analysis found that a lower cutoff point of HbA1c at 5.0% would be most efficient for diagnosing both prediabetes and diabetes in a population, and that using 5.7% was most cost-effective for undiagnosed diabetes alone.[28]

The Oregon EPC systematic review found no serious adverse effects of diabetes screening.[12] The USPSTF found that the short-term harms of screening for diabetes, such as anxiety, are small, but that the longer-term effects of labeling a large proportion of the adult U.S. population as abnormal are unknown.[11]

Intervention and Outcomes

Strong observational studies have found associations between HbA1c levels and cardiovascular risk and mortality.[29-31] For example, a large-scale study of men and women 45-79 years of age in the U.K. found that an increase in HbA1c of 1% was associated with a relative risk of death from any cause of 1.24% in men and of 1.28% in women.[29] These relative risks were independent of age, body mass index, waist-to-hip ratio, systolic blood pressure, serum cholesterol concentration, cigarette smoking, and history of cardiovascular disease. Risks for cardiovascular disease and mortality increased continuously with HbA1c concentrations through the study sample distribution. For an increasing number of stakeholders, the recognition of HbA1c as an important biomarker for cardiovascular risk strengthens the interest

and potential for use of HbA1c as a screening test for diabetics, particularly since diabetics are at higher risk for coronary heart disease and stroke.[29-31] [32] However, prospective clinical trials, especially RCTs, would be needed to establish direct, causal links between HbA1c screening and reductions in incidence of cardiovascular disease as well as diabetes.

HbA1c also provides value as an effective initial screening test for identifying individuals in whom further investigation would be useful. Rather than using fasting tests to screen all patients, HbA1c can be used for ruling in patients with prediabetic and diabetic values that require a fasting test (FPG or OGTT) to confirm a diagnosis, while ruling out patients with normal values, sparing them the discomfort of a fasting test.[33, 34] The findings of one study suggested that, with HbA1c as the initial screening test, a measurement of 5.8% could serve as the value that leads to the second fasting glucose test (FPG or OGTT).[34]   A confirmed diagnosis would prompt appropriate lifestyle or therapeutic interventions for reducing incidence of type 2 diabetes and complications. As an additional benefit, HbA1c testing can also improve the efficiency of screening since test results are not affected by short-term lifestyle and dietary changes as is the case with FPG and OGTT, but accurately reflect longer-term glycemia.[35]

In fact, there is significant evidence on the positive effect of lifestyle interventions on incidence of type 2 diabetes among individuals with prediabetes.[12] At a 3-year follow up in the Diabetes Prevention Trial (DPT), incidence of diabetes was reduced by 58% with an intensive lifestyle intervention and reduced by 31% with therapeutic intervention (i.e., metformin) in individuals determined to have IGT or IFG.[8] The STOP-NIDDM (Study to Prevent Non-Insulin Dependent Diabetes Mellitus) trial of patients with IGT showed an absolute cardiovascular risk reduction of 2.5% (although the attrition rate was 24%).[36] These studies show that that, in people known to have diabetes, certain treatment can improve health outcomes, supporting the chain of indirect evidence linking HbA1c to improved health outcomes. In contrast, the DREAM (Diabetes Reduction Assessment with Ramipril and Rosiglitazone Medication) trial, reported no significant difference in cardiovascular events between treatment and placebo groups.[37]

Additional studies of prediabetes screening followed by lifestyle interventions demonstrated effectiveness in decreasing risk and incidence of diabetes. The Finnish Diabetes Prevention study evaluated overweight adults (mean age 55 years) with IGT who were at risk for developing diabetes.[7] They were randomized to an intervention group that received individualized counseling on weight loss and related lifestyle interventions and a control group that did not. An annual OGTT test was used to identify those who progressed to diabetes--the primary endpoint. The study found that, in people with known IGT, weight loss intervention helped to reduce incidence of type 2 diabetes by 58%. Although the study did not investigate the direct impact of screening on outcomes per se, the results of this trial add further to the chain of evidence supporting clinical benefits downstream of screening. An earlier Chinese study of lifestyle interventions in people with impaired glucose tolerance also reported significant decrease (31-46%) in incidence of type 2 diabetes at a 6-year follow up.[38]

Modeling studies add to the evidence base for the value of screening for diabetes and prediabetes. For instance, a 2007 systematic review conducted in the UK evaluated screening for type 2 diabetes and IGT. Despite variable quality, structure, and assumptions of the models, the literature provided a strong case for screening for undiagnosed diabetes, including that it appears to be cost-effective for ages 40-70, especially for the older ages in this group and for

those who are hypertensive or obese. The evidence showed that detection of lesser degrees of glucose intolerance such as IGT is worthwhile because the risk of cardiovascular disease can be reduced by newer treatments to lower cholesterol levels and blood pressure and because progression to diabetes can be prevented in some of these patients. The investigators also concluded that HbA1c testing offers a good compromise between the accurate, yet more expensive and inconvenient OGTT and the FPG that can miss people with IGT. The review noted the absence of high-quality RCTs showing that screening for diabetes or IGT can reduce mortality or morbidity, but that such an RCT is underway (as described below).[14]

The findings of the first large, multicenter, international 5-year study, the Anglo-Danish-Dutch study of intensive treatment in people with type 2 diabetes detected by screening (ADDITION study) are to be published in 2010. It is anticipated that the results will provide important data on a spectrum of outcomes including CVD risk profiles, psychological status, metabolic status, and costs.[39]

There are certain practical and ethical limitations to conducting RCTs on screening and diagnostic tests. As noted by the Oregon EPC, "It is unlikely that good-quality trial evidence of the final health benefits of early glycemic control in a screening-detected population will ever be available because withholding treatment from persons with known diabetes is unethical and the length of follow-up required might be prohibitive."[11]

Several authoritative groups have issued recommendations on screening for diabetes as highlighted in Table 1.

Table 1. Screening recommendations for type 2 diabetes

Organization	Screening Recommendations
American Diabetes Association[40]	• Recommends consideration of screening to detect prediabetes (IFG or IGT) or diabetes in persons 45 years of age and older, particularly in those with a body mass index of 25 kg/m^2 or greater • Testing also should be considered in people <45 years of age and overweight if they have another risk factor[e] for diabetes • It is expected that ADA will recommend use of HbA1c as a value for diabetes and prediabetes screening in 2010
American Academy of Family Physicians[41]	• Recommends screening for type 2 diabetes in asymptomatic adults with sustained hypertension (treated or untreated) greater than 135/80 mm Hg • Concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening for type 2 diabetes in asymptomatic adults with blood pressure of 135/80 mm Hg or lower
American College of Obstetricians and Gynecology[42]	• Recommends fasting glucose testing for women beginning at age 45 years, with an interval of 3 years

---

[e] Additional risk factors for diabetes include inactivity, family history of type 2 diabetes mellitus, membership in a high-risk ethnic group, gestational diabetes, hypertension, dyslipidemia, IGT or IFG, or a history of vascular disease.

Organization	Screening Recommendations
USPSTF[11]	• Recommends screening for type 2 diabetes in asymptomatic adults with sustained blood pressure (either treated or untreated) greater than 135/80 mm Hg (grade B[f] recommendation) • Found current evidence insufficient to recommend screening for type 2 diabetes in asymptomatic adults with blood pressure 135/80 mm Hg or lower • Change in the 2008 recommendation from the 2003 recommendation is that routine screening for type 2 diabetes in adults with hyperlipidemia is no longer part of the grade B recommendation[g]
Canadian Task Force on Preventive Health Care[43]	• Recommends screening adults with hypertension and/or hyperlipidemia for type 2 diabetes mellitus to prevent cardiovascular events and death (a grade B recommendation).

Economic Value of Prediabetes Screening

Although the unit cost of HbA1c testing is currently greater than that of FPG testing, the additional benefits of HbA1c in predicting costly clinical complications support its cost-effectiveness.[17] Economic value of HbA1c testing can be realized through:

> • Reduced costs due to early detection and intervention
> • Reduced costs for treatment and management of type 2 diabetes

There have been no cost-effectiveness analyses of screening for diabetes and prediabetes with HbA1c in the context of prospective clinical trials. There is a small body of literature on cost-effectiveness analyses of screening for diabetes and prediabetes, including a few involving HbA1c testing, that use economic modeling and draw data from available clinical trials, observational studies, and various data sets.

An important earlier study (1998) evaluated the health economic impact of changes in glycemic control and quality of life on costs associated with work loss and absenteeism, functional status restrictions in daily activities, and use of health care services. A double-blind, placebo-controlled RCT assessed 569 adults with type 2 diabetes over a 15-week period. By week 15, improved glycemic control for patient in the active therapy group compared to the placebo group resulted in higher employee retention (97% vs. 85%) and retained greater productivity capacity (99% vs. 87%). There were also statistically significant differences between the groups in the changes in absenteeism, bed-days, and days of restricted activity. Absenteeism rose 8.1% for the placebo group and decreased 0.8% for the active therapy group. Reported bed days (half day or more) increased 4.4% in the placebo group and decreased 0.4% in the active therapy group. Rates of restricted activity (half day or more) also were more favorable for active therapy group. These differences demonstrated the beneficial impact of glycemic control on employee productivity.[44]

---

[f] For grade B recommendations, USPSTF recommends the service. USPSTF concludes that there is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit in moderate or substantial.

[g] USPSTF does note that clinicians should assess overall CVD risk in patients, and that if the patient's risk is near a threshold for treatment with lipid-lowering drugs, they should screen for diabetes to assess the patient's CVD risk.

A recent (2008) cost-effectiveness analysis compared four potential screening strategies with subsequent interventions for the prevention and treatment of type 2 diabetes in a hypothetical above-average risk population age 45 at the time of screening in the UK. The four strategies were: (1) screening for type 2 diabetes to enable early detection and treatment, (2) screening for type 2 diabetes and IGT, intervening with lifestyle modification in those with a diagnosis of IGT to delay or prevent diabetes, and (3) same as (2) but with pharmacological interventions, and (4) no screening. The model accounted for a 50-year follow-up of the patient population and used a 3.5% annual discount rate for costs and benefits. Screening costs included the costs of an initial FPG and a confirmatory OGTT in those who tested positive, as well as the cost of nurse time of 5 minutes for the FPG and 25 minutes for the OGTT. The investigators used sensitivity analyses to identify which model inputs had most impact on the results. Compared to no screening, the cost-effectiveness ratios of each strategy were: screening for type 2 diabetes: £14,150 ($27,860) per quality-adjusted life-year (QALY); screening for diabetes and IGT followed by lifestyle interventions: £6,242 ($12,290) per QALY; and screening for diabetes and IGT followed by pharmacological interventions: £7,023 ($13,830) per QALY. Based on the sensitivity analyses, a decision-maker who is willing to pay £20,000 ($39,400) per QALY toward each intervention, the probability of the intervention being cost effective was 49%, 93%, and 85% for each of the active screening strategies, respectively. The authors concluded that screening for type 2 diabetes and IGT in an above-average risk population age 45, with appropriate intervention for those with IGT, appears to be cost-effective, while the cost-effectiveness of a policy of screening for diabetes with no intervention to those with IGT, is still uncertain.[45]

A cost analysis published by the CDC in 2003 evaluated alternative strategies to identify individuals age 45-74 years with prediabetes (i.e., either IGT or IFG). It used data from the National Health and Nutrition Examination Survey (NHANES), 2000 census, Medicare files, and other sources. The analysis compared the effectiveness (proportion of cases identified), total costs, and efficiency (cost per case identified) of five detection strategies: OGTT, FPG, HbA1c, CBG, and a risk assessment questionnaire. For the first strategy, all individuals received an OGTT. For the other four strategies, only those with a positive screening test received an OGTT. The proportion of prediabetes and undiagnosed diabetes cases identified by each test ranged from 69% to 100%. The cost per case identified ranged from $176 to $236 from a single-payer perspective and from $247 to $332 from a societal perspective. Screening among overweight and obese individuals had a lower cost per case due to the higher prevalence of prediabetes and diabetes in those groups. Testing all with OGTT was the most effective strategy, but the CBG test and risk assessment questionnaire were the most efficient. Under an assumption that people are substantially less willing to take an OGTT than a FPG test, the FPG testing strategy was found to be the most effective strategy. The analysis quantified the tradeoffs between effectiveness and efficiency of the alternative strategies, i.e., whether the goal of the screening program is to identify more cases or to pursue the lowest cost per case. While analysis indicated higher costs for HbA1c testing (including the costs of OGTT for individuals who tested positive), the investigators noted the tradeoffs of the higher cost of HbA1c testing with that of patient nonadherence to fasting requirements for OGTT and FPG tests. CDC estimated patient nonadherence at 50% for OGTT and 75% for FPG.[46]

Another modeling study conducted in 2007 estimated the cost-effectiveness of two prediabetes screening and treatment strategies compared to no screening for overweight and obese adults age 45-74 years: (1) screening subjects and giving them the lifestyle intervention included in the Diabetes Prevention Program if they were diagnosed with both IGT and IFG and (2) screening followed by lifestyle intervention for subjects diagnosed with either IGT or IFG or both. The model assumed one-time opportunistic screening tests occurring during a scheduled physician office visit using random CPG. Positive CPGs were followed by a diagnostic OGTT or FPG; if the first diagnostic test was positive, a second was performed as confirmation. For individuals with both IGT and IFG, prediabetes screening and preventive lifestyle interventions decreased subsequent development of diabetes from 76.4% to 58.6% compared to no screening, yielding a cost-effectiveness ratio of $8,181 per QALY. Similarly, for individuals with either IGT or IFG, screening and intervention decreased subsequent development of diabetes from 57.4% to 45.2%, with a cost-effectiveness ratio of $9,511 per QALY. Sensitivity analyses showed that test attributes, including the type of screening test, its cost, sensitivity, and specificity, have small effects on the strategies' cost-effectiveness compared to the impact of the costs of lifestyle interventions. The investigators concluded that screening for prediabetes in the overweight and obese U.S. population followed by the Diabetes Prevention Program lifestyle intervention has a relatively attractive cost-effectiveness ratio.[47]

Two European modeling analyses examined the cost-effectiveness of screening using HbA1c; one in the U.K. and one in Germany. The U.K. National Health Service published a literature review and economic modeling study in 2007 of screening for type 2 diabetes using HbA1c compared to no screening or treatment until clinical detection. The economic model assessed the cost-effectiveness of a single screening round of a population with an age of 40–70 years in order to identify and treat diabetes before clinical diagnosis would occur. In addition to the 40-70 year group, the model assessed age subgroups and risk factor subgroups. The model had a 40-year time horizon. Because the literature review found that initial screening with HbA1c followed by OGTT identified more true cases than other methods (FPG or CBG), the economic model used a single HbA1c measurement as the screening test, followed by a single OGTT when the HbA1c results were > 5.7%. The model found that screening was cost-effective across the overall population age 40-70 years, at £2,266 per QALY and the subgroup populations, including 40-49 years at £10,216 per QALY, 50-59 years at £2,324 per QALY, 60-69 years at £1,152 per QALY, hypertensive patients at £1,505 per QALY, and obese patients at £1,046 per QALY, all of which are favorable cost-effectiveness ratios. The investigators noted that the costs of screening are offset in many groups by lower future treatment costs and that the cost-effectiveness of screening is determined as much by, if not more than, assumptions about the degree of control of blood glucose and future treatment protocols than by assumptions about the screening programs. They also noted that the low cost of generic statins makes an important contribution to the cost-effectiveness of diabetes screening.[14]

A German study published in 2004 compared the cost-effectiveness of four screening strategies using population-based data for those age 55-74 years. The model, which had a one-year time horizon, examined these screening strategies: FPG only, FPG + OGTT, OGTT only, and OGTT if HbA1c >5.6% (HbA1c + OGTT). OGTTs for that population yielded the lowest total costs from the payer perspective while FPG + OGTT yielded lowest costs from the societal

perspective. HbA1c + OGTT was the most expensive strategy but also the most effective in detecting new cases (54%) compared to FPG (20.7%), FPG + OGTT (25.8%), and OGTT (30%).

The effectiveness of HbA1c testing was attributed largely to the assumption of complete participation of all subjects in HbA1c testing, which can be performed as part of another scheduled physician office visit. The investigators concluded that the decision regarding which is the most favorable strategy depends on whether the goal is to identify a high number of cases or to incur lower costs at reasonable effectiveness.[48]

As a group, the available cost-effectiveness modeling studies suggest that screening for diabetes and prediabetes with follow-up treatment as appropriate can yield cost-effectiveness ratios in acceptable ranges, particularly for persons who are hypertensive or obese. However, the absence of long-term follow-up data on the cost impact of diabetes screening and treatment, limits the usefulness of current cost-effectiveness estimates. The Oregon EPC systematic review noted that the cost-effectiveness of diabetes screening programs is largely subject to the long-term health benefits of these interventions rather than the shorter-term cost of detection and treatment of diabetes, and that research on such long-term outcomes is lacking.

Conclusions

The evidence presented in this case study supports the considerable value of HbA1c testing in screening and diagnosis for diabetes and prediabetes. The high association between HbA1c levels and cardiovascular risk and mortality makes it an important biomarker for diabetic and other high-risk individuals. HbA1c has comparable diagnostic accuracy and predictive value to other tests currently used as the standard of care, and improves consistency in testing as its test results are not affected by short-term glycemic fluctuations (due to current effects of diet, medications, and other factors) that affect other tests. Many studies have confirmed the effectiveness of HbA1c testing in screening for identifying prediabetic individuals who will benefit from early interventions, contributing to substantial decreases in incidence of type 2 diabetes. Use of HbA1c increases patient-centeredness of care given the improved convenience and comfort of testing for the patient. Findings from cost-effectiveness modeling of various testing modalities for diabetes and prediabetes indicates that screening can be cost-effective, especially for particular at-risk groups. Although HbA1c testing (typically used in combination with confirmatory OGTT) tends to be more costly than alternative types of testing, it is able to detect more cases of diabetes in a screened population and yields cost-effectiveness ratios in favorable ranges. Still, the cost-effectiveness of screening programs for diabetes and prediabetes awaits further data from longer-term follow-up studies.

References

1. Total prevalence of diabetes and pre-diabetes. American Diabetes Association, 2009. (Accessed April 17, 2009, at http://www.diabetes.org/diabetes-statistics/prevalence.jsp.)

2. National diabetes fact sheet, 2007. Centers for Disease Control and Prevention, 2007. (Accessed May 1, 2009, at http://www.cdc.gov/diabetes/pubs/factsheet07.htm.)

3. Direct and indirect costs of diabetes in the United States. American Diabetes Association, 2008. (Accessed June 26, 2008, at http://www.diabetes.org/diabetes-statistics/cost-of-diabetes-in-us.jsp.)

4. Goetzel R.Z., Long SR, Ozminkowski RJ, et al. Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting U.S. employers. J Occup Environ Med 2004;46:398.

5. Harris MI. Undiagnosed NIDDM: clinical and public health issues. Diabetes Care 1993;16:642-52.

6. Gulliford MC, Charlton J, Latinovic R. Increased utilization of primary care 5 years before diagnosis of type 2 diabetes. Diabetes Care 2005;28:47-52.

7. Tuomilehto J, Lindstrom J, Eriksson JG, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001;344:1343-50.

8. Knowler WC, Barrett-Connor E, Fowler SE, et al. Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002;346(6):393-403.

9. The Diabetes Prevention Program Research Group. The Diabetes Prevention Program. Design and methods for a clinical trial in the prevention of type 2 diabetes. Diabetes Care 1999;22:623-34.

10. The Diabetes Prevention Program Research Group. The Diabetes Prevention Program. Baseline characteristics of the randomized cohort. Diabetes Care 2009;23:1619-29.

11. U.S. Preventive Services Task Force. Screening for type 2 diabetes mellitus in adults: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2008;148(11):846-54.

12. Norris SL, Kansagara D, Bougatsos C, Fu R. Screening adults for type 2 diabetes: a review of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2008;148:855-68.

13. Lewandrowski K, Lee-Lewandrowski E. Capillary blood glucose testing at the point of care--clinical applications and the evolution of diagnostic technologies. Business Briefing: Global HealthCare Issue 2002;3:76-9.

14. Waugh N, Scotland G, McNamee P, et al. Screening for type 2 diabetes: literature review and economic modeling. Health Technol Assess 2007;11(7).

15. Expert panel proposes HbA1c replace glucose levels for diabetes. Modern Medicine, 2009. http://www.modernmedicine.com/modernmedicine/Modern+Medicine+Conference+N

ews/ADA-2009-Expert-committee-proposes-HbAsub1csub-rep/ArticleStandard/Article/detail/602012.)

16. International Expert Committee. International Expert Committee report on the role of the HbA1c assay in the diagnosis of diabetes. Diabetes Care 2009;32(7):1-8.

17. Bennett CM, Guo M, Dharmage SC. HbA1C as a screening tool for detection of type 2 diabetes: a systematic review. Diabet Med 2007;24:333-43.

18. Wang W, Lee ET, Fabsitz R, et al. Using HbA(1c) to improve efficacy of the American Diabetes Association fasting plasma glucose criterion in screening for new type 2 diabetes in American Indians: The Strong Heart Study. Diabetes Care 2002;25:1365-70.

19. Ito C, Maeda R, Ishida S, et al. Correlation among fasting plasma glucose, two-hour plasma glucose levels in OGTT and HbA1c. Diabetes Res Clin Pract 2000;50:225-30.

20. Jesudason DR, Dunstan K, Leong D, Wittert GA. Macrovascular risk and diagnostic criteria for type 2 diabetes: implications for the use of FPG and HbA(1c) for cost-effective screening. Diabetes Care 2003;26:485-90.

21. Tanaka Y, Atsumi Y, Matsuoka K, et al. Usefulness of stable HbA(1c) for supportive marker to diagnose diabetes mellitus in Japanese subjects. Diabetes Res Clin Pract 2001;53:41-5.

22. Tavintharan S, Chew LS, Heng DM. A rational alternative for diagnosis of diabetes mellitus in high-risk individuals. Ann Acad Med Singapore 2000;29:213-8.

23. Herdzik E, Safranow K, Ciechanowski K. Diagnostic value of fasting capillary glucose, fructosamine and glycosylated haemoglobin in detecting diabetes and other glucose tolerance abnormalities compared to oral glucose tolerance test. Acta Diabetol 2002;39:15-22.

24. Ko GT, Chan JC, Yeung VT, et al. Combined use of a fasting plasma glucose concentration and HbA1c for fructosamine predicts the likelihood of having diabetes in high-risk subjects. Diabetes Care 1998;21:1221-5.

25. Wiener K, Roberts NB. The relative merits of haemoglobin A1c and fasting plasma glucose as first-line diagnostic tests for diabetes mellitus in non-pregnant subjects. Diabet Med 1998;15:558-63.

26. Edelman D, Olsen MK, Dudley TK, et al. Utility of hemoglobin A1c in predicting diabetes risk. J Gen Intern Med 2004;19(12):1175-80.

27. Norberg M, Eriksson JW, Lindahl B, et al. A combination of HbA1c, fasting glucose and BMI is effective in screening for individuals at risk of future type 2 diabetes: OGTT is not needed. J Intern Med 2006;260(3):263-71.

28. Zhang P, Engelgau MM, Valdez R, et al. Efficient cutoff points for three screening tests for detecting undiagnosed diabetes and prediabetes. Diabetes Care 2005;28(6):1321-5.

29. Khaw KT, Wareham N, Bingham S, et al. Association of hemoglobin A1c with cardiovascular disease and mortality in adults: the European prospective investigation into cancer in Norfolk. Ann Intern Med 2004;141:413-20.

30. Gerstein HC. Glycosylated hemoglobin: finally ready for prime time as a cardiovascular risk factor. Ann Intern Med 2004;141:475-6.

31. Selvin E, Marinopoulos S, Berkenblit G, et al. Meta-analysis: glycosylated hemoglobin and cardiovascular disease in diabetes mellitus. Ann Intern Med 2004;141:421-31.

32. Saudek CD, Herman WH, Sacks DB, Bergenstal RM, Edelman D, Davidson MB. A new look at screening and diagnosing diabetes mellitus. J Clin Endocrinol Metab 2008;93:2447-53.

33. Rohlfing CL, Wiedmeyer HM, Little RR, et al. Defining the relationship between plasma glucose and HbA(1c): analysis of glucose profiles and HbA(1c) in the Diabetes Control and Complications Trial. Diabetes Care 2002;25(2):275-8.

34. Buell C, Kermah D, Davidson MB. Utility of A1C for diabetes screening in the 1999-2004 NHANES population. Diabetes Care 2007;30(9):2233-5.

35. Tahara Y, Shima K. The response of GHb to stepwise plasma glucose change over time in diabetic patients. Diabetes Care 1993;16(9):1313-4.

36. Chiasson JL, Josse RG, Gomis R, et al. Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomized trial. Lancet 2002;359:2072-7.

37. DREAM Trial Investigators, Gerstein HC, Yusuf S, Bosch J, et al. Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomized controlled trial. Lancet 2006;368:1096-105.

38. Pan XR, Li GW, Hu YH, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study. Diabetes Care 1997;20(4):537-44.

39. Lauritzen T, Griffin S, Borch-Johnsen K, et al. The ADDITION study: proposed trial of the cost-effectiveness of an intensive multifactorial intervention on morbidity and mortality among people with type 2 diabetes detected by screening. Int J Obes Relat Metab Disord 2000;24(Supplement 3):S6-S11.

40. American Diabetes Association. Standards of medical care in diabetes--2009. Diabetes Care 2009;32(1):S13-S61.

41. Summary of recommendations for clinical preventive services. Revision: 6.8, April 2009. American Association of Family Physicians, 2009. (Accessed August 20, 2009, at http://www.aafp.org/online/etc/medialib/aafp_org/documents/clinical/CPS/rcps08-2005.Par.0001.File.tmp/CPSapprovedApril2009.pdf.)

42. Guidelines for women's health care. A resource manual. Report No. Third Edition. Washington, DC: American College of Obstetricians and Gynecologists, 2007.

43. Feig DS, Palda VA, Lipscombe L, with Canadian Task Force on Preventive Health Care. Screening for type 2 diabetes mellitus to prevent vascular complications: updated recommendations from the Canadian Task Force on Preventive Health Care. Can Med Assoc J 2005;172(2):177-80.

44. Testa MA, Simonson DC. Health economic benefits and quality of life during improved glycemic control in patients with type 2 diabetes mellitus: a randomized, controlled, double-blinded trial. JAMA 1998;280(17):1490-6.

45. Gillies CL, Lambert PC, Abrams KR, et al. Different strategies for screening and prevention of type 2 diabetes in adults: cost effectiveness analysis. Br Med J 2008;336(7654):1180-5.

46. Zhang P, Engelgau MM, Valdez R. Costs of screening for pre-diabetes among U.S. adults: a comparison of different screening strategies. Diabetes Care 2003;26(9):2536-42.

47. Hoerger TJ, Hicks KA, Sorensen SW, et al. Cost-effectiveness of screening for pre-diabetes among overweight and obese U.S. adults. Diabetes Care 2007;30(11):2874-9.

48. Icks A, Haastert B, Gandjour A, et al. Cost-effectiveness analysis of different screening procedures for type 2 diabetes: the KORA survey 2000. Diabetes Care 2004;27(9):2120-8.

Case Study III:

Diagnostic Testing for KRAS Genetic Mutation in those with Metastatic Colorectal Cancer

Magnitude and Importance of Colorectal Cancer

Colorectal cancer includes cancers of the colon, rectum, and appendix. In 2009, there will be an estimated 148,000 newly diagnosed cases and close to 50,000 deaths from colorectal cancer. For men and women, colorectal cancer ranks third in incidence and third in cause of cancer death. Incidence of colorectal cancer has decreased during the last two decades, from 66.3 cases per 100,000 in 1985 to 46.4 in 2005. Increased screening contributed partly to these improvements, as screening tests can lead to detection and removal of polyps before they progress to cancer, thereby decreasing mortality, morbidity, and costs of care.[1]

According to the American Cancer Society (ACS), the 1-, 5-, and 10-year relative survival rates for persons diagnosed with colorectal cancer are 83%, 64%, and 58%, respectively, although survival may differ by stage of disease. For those patients in whom colorectal cancers are detected at an early, localized stage, the 5-year survival is 90%; however, only 40% of colorectal cancers are diagnosed at this stage. About 25% of patients with colorectal cancer present with overt symptoms of metastatic disease and 40-50% of those who are newly diagnosed develop metastatic disease.[2] Once the cancer has spreads regionally to involve adjacent organs or lymph nodes in these patients, the 5-year survival drops to 68%. For persons with distant metastases, 5-year survival is only 11%.[1]

There are no current, rigorous estimates of the total economic burden of colorectal cancer.[3] Such an estimate would include costs associated with screening, surveillance, diagnosis, hospitalization, surgery, radiotherapy, anticancer agents, supportive care, physician charges, clinic visits, laboratory fees, and medications.[4] The highest costs of treatment are incurred during the early stage of the disease (i.e., surgery, surveillance, monitoring) and during the terminal stage (i.e., hospitalization, chemo-and immunotherapy, and supportive care). Based on older data, a 2003 publication projected that the combined inpatient and outpatient costs of colorectal cancer in the US were in the range of $5.5-6.5 billion.[5] Extrapolated to 2009, this would amount to approximately $7.5-9.5 billion; however, this very rough estimate does not reflect changes in costs associated with colorectal cancer incidence, detection, or management in recent years.

New Treatments for Colorectal Cancer

The main types of treatment for colorectal cancer are surgery, radiation, chemotherapy, and, for certain cancers, monoclonal antibody (MAb) therapies and vaccines. Treatment choice depends on the stage of cancer and other factors, and patients often receive combinations of these treatments. Surgery is the most common treatment for colorectal cancer and can be curative for non-metastatic cancers. Radiation also may be given depending on the type of tumor. Chemotherapy is typically given to those with more advanced cancers.

Three MAb therapies that were recently approved by FDA target specific sites on cancer cells. Bevacizumab (Avastin) blocks the growth of blood vessels to the tumor. Cetuximab (Eribitux)

and panitumumab (Vectibix) block the effect of hormone-like factors that promote cancer cell growth.[1, 5, 6] The targets of cetuximab and panitumumab are epidermal growth factor receptors (EGFRs), which are signaling proteins on cells that normally control cell division. EGFR inhibitors such as cetuximab and panitumumab bind to EGFRs, thereby interfering with EGFR-mediated cell proliferation.

The effectiveness of MAb therapies in colorectal cancer can vary depending on the genetic makeup of the tumor. In some patients, the tumor cells have a mutation in a gene known as *KRAS (v-Ki-ras2 Kirsten rat sarcoma viral oncogene)* and may cause a tumor to be non-responsive to cetuximab and panitumumab.[a] As such, *KRAS* status is an important predictor of non-response to targeted therapy in patients with metastatic colorectal cancer. The normal or non-mutant form of *KRAS* is also knows as the *wild-type KRAS*. Since the use of EGFR inhibitors frequently results in skin rashes and other adverse side effects, determination of *KRAS* status can help to avoid ineffective treatment and unnecessary exposure to any accompanying adverse effects, and redirect treatment decisions to alternative therapies.

Laboratory Tests for Treatment of Colorectal Cancer

Molecular diagnostic tests are increasingly important tools for understanding individual variation in response to drug therapies, including their benefits and adverse effects. As in the case of EGFR inhibitors, patient response to these therapies can be influenced by genetic variability in drug targets (e.g., receptors on cancer cells) or drug disposition (e.g., drug uptake, transport, and metabolism).[7, 8] The use of individual genetic or genomic information to select the most effective treatment or to avoid those with severe adverse effects for that individual is known as "pharmacogenomics."

Diagnostic colonoscopies provide visual confirmation of colorectal cancer and enable removal of polyps and other tissue samples for pathology testing to determine the stage of disease and any tumor characteristics that might inform treatment decisions. During surgical resections, tumor samples are sent to the laboratory for molecular testing to identify molecular-level alterations in the tumor cells that inactivate tumor-suppressor genes or activate oncogenes.[9] Gene mutations that can trigger development and progression of metastatic colorectal cancer include those in the genes *KRAS, tumor protein p53,* and *adenomatous polyposis coli (APC)*. Mutation of the *KRAS* gene is the most common oncogene and a significant prognostic factor for treatment effectiveness.

Polymerase chain reaction (PCR) is one molecular-based technique used in *KRAS* mutation testing to distinguish patients whose tumors have the *KRAS* mutation from patients whose tumors have the (normal) *wild-type KRAS*. As described below, only those patients with wild-type tumors have the potential to respond to anti-EGFR antibody therapy.

---

[a] The *KRAS* gene encodes the protein KRAS, which acts as an on/off switch, transmitting growth signals from EGFRs on cell membranes. Mutated *KRAS* genes are potent oncogenes that play a role in many cancers, including leukemias and lung, pancreatic, and colorectal cancer.

Clinical Value of *KRAS* Mutation Screening

The clinical analytic framework presented in Figure 1 depicts the clinical and economic value of *KRAS* mutation testing to patients with metastatic colorectal cancer. *KRAS* mutation testing can improve clinical value in several different ways, including:

- Improves capacity to identify individuals with *KRAS* gene mutations that will not respond to EGFR inhibitors

- Provides an evidence-based way to select treatment that is more likely to be effective

- Decreases patient exposure to adverse effects that would have accompanied an ineffective treatment

- Leads to improved patient outcomes, including improved decreased morbidity, improved survival, and improved quality of life

Detection

KRAS testing is regulated under the Clinical Laboratory Improvement Amendments (CLIA) and are available from several major laboratories. Premarket approval from FDA is not required when the testing is performed in a laboratory that is licensed by CLIA for high-complexity testing.[16] The current methods of *KRAS* testing have not been standardized, with perhaps 20 different methods of *KRAS* genotyping. Many of these methods were developed in-house by laboratories, i.e., laboratory-developed tests or for research purposes only. As a result, there can be some variation in the degree of analytic validation among the different methods, and no rigorous, well-designed studies have been done to compare the diagnostic accuracy of different techniques.

To assess the sensitivity and specificity of *KRAS* mutation testing, investigators recently conducted a meta-analysis of eight available studies that met certain methodological inclusion criteria. As a group, the eight studies accounted for a total of 817 patients, of which 306 patients have mutated *KRAS*. The eight studies included some involving EGFRs alone and some in combination with chemotherapy. In this instance, sensitivity referred to the ability of the *KRAS* mutation test to identify patients with *KRAS* mutations showing no response to EGFR inhibitors. Specificity referred to the ability of the *KRAS* mutation test to identify patients with wild-type *KRAS* showing a response to EGFR inhibitors.[10]

The investigators found that the pooled (combined across studies) specificity of *KRAS* mutation testing was high (pooled estimate of 93%, 95% confidence interval [CI] 93-97%). This suggests that a response to EGFR inhibitors is highly unlikely in the presence of a *KRAS* mutation, supporting the use of *KRAS* mutation testing to rule out treatment with EGFR inhibitors for those patients who are found to have the *KRAS* mutation. However, the investigators found that the pooled sensitivity of *KRAS* mutation testing (pooled estimate of 47%, 95% CI 43-52%) was relatively low. This suggests that a positive test for *KRAS* mutation fails to identify a large proportion of non-responders, i.e., that non-response to EGFR inhibitors is not confined to patients with *KRAS* mutations, but also occurs in a substantial number of patients with *wild-type KRAS*. The investigators concluded that *KRAS* mutations are highly specific negative predictors of response to EGFR-inhibitors alone or in combination with chemotherapy in patients with metastatic colorectal cancer.

Figure 1. Analytic Framework for *KRAS* Mutation Testing
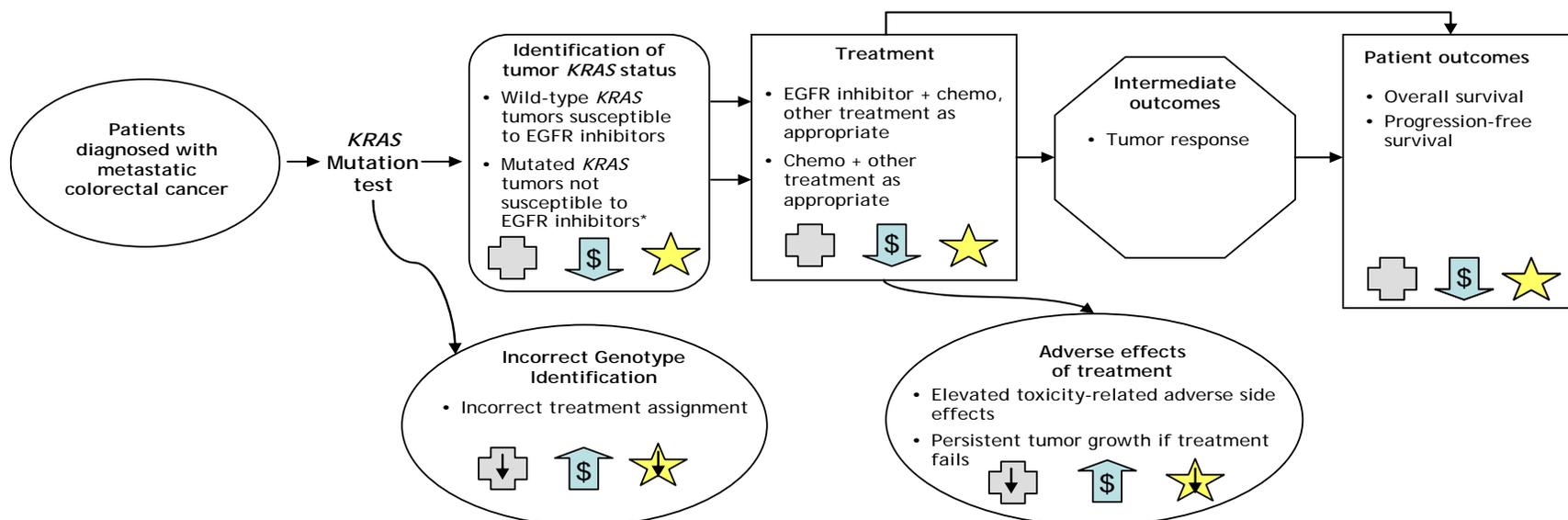
Intermediate and Patient Outcomes

Analyses of RCT data demonstrate that EGFR inhibitor therapy for patients with *wild-type KRAS* patients result in:
- Improved tumor response[1-3,5]
- Improved progression-free survival[1,3,5]
- Improved overall survival[3]

Improved Quality of Life

Analysis of RCT data demonstrate that EGFR inhibitor therapy for patients with *wild-type KRAS* patients result in:
- Less physical function deterioration at 8 weeks, improved global health status at 8 weeks, preserved global health status at 16 weeks.[8]

Cost Savings and Cost Effectiveness
- No published direct evidence of cost savings; one unpublished analysis estimates hundreds of millions of dollars per year saved by ruling out non-beneficial EGFR inhibitor treatment for patients with KRAS mutations.[6]
- Manufacturer's modeling estimates that, in patients identified with *wild-type KRAS,* ICERs under best-case scenarios are £28,024-34,646/QALY for adding cetuximab to FOLFIRI and £29,327-40,529/QALY for adding cetuximab to FOLFOX (submitted to UK NICE; other scenarios had higher ICERs).[7]

*EGFR inhibitors– epidermal growth factor receptor inhibitors: e.g., cetuximab, panitumumab

Sources:
1. Amado RG, Wolf M, Peeters M, et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. J Clin Oncol 2008;26:1626-34.
2. Van Cutsem E, Kohne CH, Hitre R, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. N Engl J Med 2009;360(14):-1408.
3. Karapetis CS, Khambata-Ford S, Jonker DJ, et al. K-ras mutations benefit from cetuximab in advance colorectal cancer. N Engl J Med 2008;359:1757-65.
4. Punt CJ, Tol J, Rodenburg CJ, et al. Randomized phase III study of capecitabine, oxaliplatin, and bevacizumab with or without cetuximab in advanced colorectal cancer (ACC), the CAIRO2 study of the Dutch Colorectal Cancer Group (DCCG). Journal of Clinical Oncology 2008;26(Supplement):abstr LBA4011.
5. Bokemeyer I, Bondarenko JT, Hartmann FG, et al. KRAS status and efficacy of first-line treatment of patients with metastatic colorectal cancer (mCRC) with FOLFOX with or without cetuximab: the OPUS experience. J Clin Oncol 2008;26(Supplement):4000.
6. Van Cutsem E, Lang I, D'haens G, et al. KRAS status and efficacy in the first-line treatment of patients with metastatic colorectal cancer (mCRC) treated with FOLFIRI with or without cetuximab: the CRYSTAL experience. J Clin Oncol 2008;26(Supplement):2.
7. Jonker DJ, Karapetis CS, Moore M, et al. Randomized phase III trial of cetuximab monotherapy plus best supportive care (BSC) versus BSC alone in patients with pretreated metastatic epidermal growth factor receptor-positive colorectal carcinoma: a trial of the National Cancer Institute of Canada Clinical Trials Group and the Australasian Gastrointestinals Trials Group. American Association for Cancer Research Annual Meeting, 2007.
8. NICE single technology appraisal of cetuximab for the first line treatment of metastatic colorectal cancer. U.K. National Institute for Health and Clinical Excellence, 2009.

Further research is needed to improve identification of patients who are unlikely to respond to EGFR inhibitors. The relatively low sensitivity of testing for *KRAS* mutations for determining non-responsiveness strongly suggest that there are additional mechanisms of resistance to EGFR inhibitors.[10] Indeed, *KRAS* gene mutations appear to account for only about 35-45% of patients who are non-responders to EGFR inhibitors. Some investigators suggest that a higher percentage of non-responsive patients can be identified when other molecular alterations in the EGFR signaling pathway are analyzed, including involving such genes known as *PIK3CA*, *BRAF*, and *PTEN*, although these remain to be validated for this purpose.[11]

The high predictive value of *KRAS* testing to identify lack of response to EGFR inhibitors in patients with metastatic colorectal cancer tumors also suggests that *KRAS* mutation status may be predictive of response in certain other cancers with high prevalence of *KRAS* mutations.[12, 13] However, further research is needed to identify and validate potential roles for *KRAS* testing.

Intervention and Outcome

A recent set of retrospective analyses of RCT data demonstrate that the response of patients with metastatic colorectal cancer to cetuximab or panitumumab is strongly influenced by their *KRAS* status. As such, the value of *KRAS* testing is based on its ability to distinguish between those patients who are highly unlikely to respond to EGFR inhibitors, and therefore can be directed to potentially beneficial alternative therapy, and those who are more likely to respond to EGFR inhibitors. These analyses have been compiled in reviews conducted by the Blue Cross Blue Shield Association Technology Evaluation Center (BCBS TEC),[16] American Society of Clinical Oncology (ASCO), and FDA.[31] The FDA review, which is the most recent, includes analyses of seven RCTs, two of which do not appear to have been published to date; both of the others include analyses of the other five RCTs. (BCBS TEC and ASCO also evaluated five single-arm studies for tumor response according to *KRAS* status.) Although these RCTs did not prospectively assign patients to treatment according to their *KRAS* gene status, these analyses retrospectively compared how well patients with *KRAS* mutations and *wild-type KRAS* responded to the EGFR inhibitor therapies. The outcomes assessed in these analyses included two or more of overall response rate (tumor response),[b] progression-free survival,[c] overall survival,[d] and quality of life.[14-18]

In these RCTs, patients received standard of care (i.e., basic supportive care or standard chemotherapy) and were randomized to receive an EGFR inhibitor (cetuximab or panitumumab) or no additional therapy. The percentage of the original RCT populations for which *KRAS* status was assessed ranged from 23% to 92%. Among patients whose tumors had *KRAS* mutations, those who received an EGFR inhibitor did no better or worse in nearly all patient outcomes than the control groups that did not receive an EGFR inhibitor. For example, of the eight primary outcomes measured across the seven studies (progression-free survival in four studies, tumor response in two studies, and overall survival in two studies), patients whose tumors had *KRAS* mutations receiving an EGFR inhibitor did worse than the control group in five and no better than the control group in three. As confirmed by BCBS TEC, ASCO,

---

[b] Overall response rate is an intermediate outcome that refers to the percentage of patients whose tumors were reduced in size as a result of treatment.

[c] Progression-free survival is the length of time during and after treatment when a patient's disease does not worsen. Progression-free survival does not refer to longer patient survival.

[d] Overall survival is the total amount of time that a patient survives following treatment, regardless of the cause of death.

and FDA, these analyses strongly suggest that EGFR inhibitors are not effective for treatment of patients with metastatic colorectal cancer containing *KRAS* mutations.

Among the patients in these five RCTs who received EGFR inhibitor and were found subsequently to have *wild-type KRAS,* progression-free survival was significantly better in three of the five trials in which it was reported whereas overall survival was significantly improved in only one of the three trials in which it was reported. In the latter trial which overall survival was improved, the benefit of being in the cetuximab group was modest, i.e., 9.5 months for the cetuximab group vs. 4.8 months for the best supportive group. There is limited evidence concerning impact on health-related quality of life. In an analysis of the data from one of the RCTs, patients with *wild-type KRAS* who received cetuximab experienced less physical function deterioration at 8 weeks, improved global health status at 8 weeks, and preserved global health status at 16 weeks.[19]

As noted above, although the available research demonstrates that *wild-type KRAS* status in patients with metastatic colorectal cancer is strongly associated with favorable response to cetuximab, including objective response and progression-free survival, available evidence also indicates that about half of *wild-type KRAS* patients may be non-responders.[20]

Another benefit to *KRAS* testing is the potential for avoidance of adverse effects. Among patients with advanced colorectal cancer treated with cetuximab monotherapy, the more frequently reported adverse effects include: acneform rash/desquamation (89%); fatigue (89%); abdominal pain (59%); dyspnea (48%); gastrointestinal (35-39%), infection (35%), fever (30%), stomatitis (25%), and adverse infusion reactions (20%).[1] Certainly, many of these side effects occur with other chemotherapies and MAb; however, additional incidence of adverse effects from cetuximab can be considerable. For example, in one of the analyses of RCT data noted above, incidence of treatment-related serious adverse events was 26% for the cetuximab+FOLFIRI group versus 19.3% for the FOLFIRI group. The following more severe reactions were significantly more frequent with cetuximab+FOLFIRI than with FOLFIRI alone: skin reactions (19.7% vs. 0.2%), infusion-related reactions (2.5% vs. 0%), and diarrhea (15.7% vs. 10.5%).[21] (These were grade 3 or 4 adverse events, based on the National Cancer Institute Common Toxicity Criteria, a four-point scale [0-4], with 4 indicating most severe reaction.[22]) Avoiding or minimizing these adverse effects to the extent possible through testing of *KRAS* mutation status in patients with metastatic colorectal cancer addresses the goals of patient-centered, safe, effective, quality of care supported through pharmacogenomics.[23]

Incorporation into Practice Guidelines

Given the available evidence regarding the role of *KRAS* status in predicting response to EGFR inhibitor therapies, the National Comprehensive Cancer Network (NCCN) has updated its clinical guidelines for colorectal cancer in 2008 to include the recommendation that "determining *KRAS* gene status of either the primary tumor or a site of metastasis should be part of the pre-treatment work-up for all patients with metastatic colorectal cancer." Further, the EGFR inhibitors, cetuximab and panitumumab, "either as single agents, or, in the case of cetuximab, in combination with other agents, are now recommended only for patients with tumors characterized by the *wild-type KRAS* gene," according to NCCN.[24]

In 2009, ASCO issued a provisional clinical opinion stating that "all patients with metastatic colorectal carcinoma who are candidates for anti-EGFR antibody therapy should have their

tumor tested for *KRAS* mutations in a CLIA-accredited laboratory. If *KRAS* mutation in codon 12 or 13 is detected, then patients with metastatic colorectal carcinoma should not receive anti-EGFR antibody therapy as part of their treatment."[25]

In July 2009, FDA announced revisions to the prescribing information for EGFR inhibitors and colorectal cancer, requiring inclusion of information on variations in the *KRAS* gene that may affect patient response to those therapies. The agency also updated the clinical studies section of the label to include results from retrospective analyses based on data from seven RCTs with drugs in this class.[26]

In assessing *KRAS* testing for informing decisions about EGFR inhibitor therapy for metastatic colorectal cancer, BCBS TEC found that *KRAS* testing met all five of its evaluation criteria. With regard to the clinical utility of *KRAS* gene testing, BCBS TEC stated that: "Based on clinical validity of *KRAS* mutation testing to predict non-response to EGFR inhibitors, the clinical utility can be inferred to guide treatment decisions in patients with metastatic colorectal cancer."[16]

Economic Value of *KRAS* Mutation Screening

To date, there are no published peer-reviewed cost analyses of *KRAS* mutation testing.[3, 27] The potential economic value of *KRAS* testing resides in costs savings and improved cost effectiveness of appropriately targeted EGFR inhibitors for patients with advanced, metastatic colorectal cancer. These include:

- Cost savings realized from avoiding ineffective treatment

- Cost savings from avoiding management of adverse effects of EGFR inhibitor therapy in patients who would not benefit from that therapy

- Improved cost effectiveness resulting from better use of EGFR inhibitor therapy

A few unpublished studies address the cost effectiveness of alternative therapies involving EGFR inhibitors for metastatic colorectal cancer. However, these analyses do not address the particular contribution to cost effectiveness of using *KRAS* testing to inform treatment choices.

As part of its recent appraisals of MAb therapies for first-line treatment of metastatic colorectal cancer, the UK National Institute for Clinical Excellence (NICE) reported on cost-effectiveness analyses submitted by manufacturers. Consistent with the licensed indication for cetuximab, these cost-effectiveness analyses assumed that all patients had been tested for *KRAS* mutations and that only those with *wild-type KRAS* were treated with regimens involving cetuximab. The results of these varied according to the assumptions about treatment approaches used, e.g., rules about when to stop treatment based on patient response after a certain number of weeks, time horizon (e.g., 10 years vs. lifetime), rates for liver resections (how often required and their success rates) that would be required as part of treatment, and use of a manufacturer rebate program. For example, a manufacturer's model estimated that, under base case scenarios, the incremental cost-effectiveness ratio (ICER) of treating with cetuximab+FOLFIRI[e] vs. FOLFIRI

---

[e] FOLFIRI is a standard chemotherapy regimen for treatment of colorectal cancer, that includes the drugs FOL – folinic acid (leucovorin), F – fluorouracil (5-FU), and IRI – irinotecan.

alone was in the range of £30,546-69,287/QALY, and that the ICER for cetuximab+FOLFOX[f] vs. FOLFOX alone was in the range of £37,571-63,245/QALY. Under certain best case scenarios, the ICER of treating with cetuximab+FOLFIRI vs. FOLFIRI alone was in the range of £28,024-34,646/QALY, and that the ICER for cetuximab+FOLFOX vs. FOLFOX alone was in the range of £29,327-40,529/QALY.[28, 29] As such, these analyses submitted by manufacturers indicate that these treatments involving cetuximab are borderline cost effective. (NICE cites £30,000/QALY as an approximate threshold above which it would not consider an intervention to be cost effective.)

The National Cancer Institute of Canada Clinical Trials Group CO.17 study, as reported in 2009, demonstrated higher cost-effectiveness ratios for cetuximab given in addition to best supportive care for metastatic colorectal cancer.[30] Analyses included direct medical costs: medications, physician visits, toxicity management, blood products, emergency department visits, and hospitalizations. Mean survival times were calculated for the entire population and for the subset of patients with *wild-type KRAS* tumors over an 18-19-month period. Study results are presented in Table 1.

Table 1. Clinical and economic benefit of cetuximab compared to best supportive care

	Entire population	Patients with *wild-type KRAS*
Clinical benefit		
Quality-adjusted survival	0.12 years	0.28 years
Quality-adjusted life years	0.08 QALYs	0.18 QALYs,
Cost effectiveness		
Incremental cost	$23,969	$33,617
Cost effectiveness ratio	$199,742 per life-year gained (95% CI = $125 973 to $652 492 per life-year gained)	$120,061 per life-year gained (95% CI = $88 679 to $207 075 per life-year gained)
Cost utility ratio	$299,613 per QALY gained (95% CI = $187 440 to $898 201 per QALY gained)	$186,761 per QALY gained (95% CI = $130 326 to $334 940 per QALY gained)

The study demonstrated modest improvement in overall and quality-adjusted survival (QAS) and QALYs for patients with *wild-type KRAS* who received cetuximab compared to the entire patient population.[30] As shown in Table 1, for those with *wild-type KRAS,* mean gains were QAS of 0.28 years and 0.18 QALYs versus the entire population with QAS of 0.12 years and 0.08 QALYs. The incremental cost for use of cetuximab compared with best supportive care was lower for the full patient population compared to patients with *wild-type KRAS*. However, limiting the analysis to patients with *wild-type KRAS* demonstrated lower overall costs per life-year gained for *wild-type-KRAS* versus the entire population, though both ratios were high. Similarly, the cost per QALY ratio for *wild-type KRAS* was lower than for the entire population, though both ratios were high. In a sensitivity analysis, cetuximab cost and patient survival were the only variables that influenced cost effectiveness. The investigators concluded that ICERs of

---

[f] FOLFOX is a standard chemotherapy regimen for treatment of colorectal cancer that includes the drugs FOL– folinic acid (leucovorin), F – fluorouracil (5-FU), and OX – oxaliplatin.

cetuximab over best supportive care alone in unselected advanced colorectal cancer patients are high and sensitive to drug cost.

Another recently unpublished analysis projected considerable cost savings of using *KRAS* testing to inform treatment decisions. Using 2008 ACS data, the annual incidence of metastatic colorectal cancer was estimated to be 29,762 new cases per year.[31] Upfront *KRAS* testing for all of these new cases would cost $13 million ($452 per patient). The cost of cetuximab treatment was estimated to be $71,120 per patient. Assuming that the 36% of patients found to have *KRAS* mutations would not receive cetuximab, drug cost savings were estimated at $753 million. Considering the cost of *KRAS* testing, net savings in ineffective treatment with cetuximab would be $740 million. (This estimate may be high for several reasons. It assumes that, in the absence of *KRAS* testing, all new patients would have received 24 weekly doses of cetuximab. It also assumes that, in the new scenario, all patients would receive upfront *KRAS* testing, and all patients with *KRAS* mutations would be identified and that none of these patients would receive cetuximab. Also, the analysis does not appear to account for the costs of alternative treatment that would substitute for cetuximab in the patients with *KRAS* mutations.) The investigators noted that, although cetuximab is used as a second- or third-line therapy, *KRAS*-based treatment selection is likely to produce cost savings across all lines of therapy since validated predictive molecular markers may not only spare patients ineffective and toxic therapies, but may also greatly reduce futile costs.[31]

Conclusions

KRAS testing is a current example of advances in molecular and genetic testing that are becoming important diagnostic tools for understanding individual response to treatments. The results of these tests are having a direct influence on clinical decision making, patient outcomes, and costs. Identifying patients with *KRAS*-mutated tumors is helping to avoid prescription of ineffective drugs, unnecessary exposure to adverse effects from those drugs, and more prompt treatment with alternative drug regimens.

Based on retrospective subgroup analyses of data from multiple RCTs and other clinical studies, the response of patients with metastatic colorectal cancer to cetuximab or panitumumab as monotherapies or in combination with other drugs differs considerably based on their *KRAS* gene status. Patients with *KRAS* mutations show virtually no response to these drug regimens. The benefits in tumor response, progression-free survival, and overall survival found in varying extents across available studies have been limited only to those patients with *wild-type KRAS*.

These findings demonstrate that *KRAS* gene status should be a key factor in treatment decisions for patients with metastatic colorectal cancer that may involve use of EGFR inhibitors. The role of *KRAS* testing is reflected in recently modified FDA prescribing information in product labels for EGFR inhibitors, and in clinical practice guidelines of NCCN, ASCO, and other authoritative sources calling for pretreatment work-up for all patients with metastatic colorectal cancer.

Information about the sensitivity and specificity of *KRAS* mutation testing is limited and variable. Performed as CLIA-regulated laboratory-developed tests, *KRAS* mutation testing is not standardized across laboratories, which may result in inconsistencies in results across laboratories.

Although *KRAS* testing to inform treatment decisions for these patients is strongly indicated based on available evidence, the actual impact of the testing on clinical decisions or on patient outcomes has not yet been quantified in prospective studies. This would include demonstrating that the test performs as well in upfront testing of new patients in clinical practice as it has in retrospective subgroup analyses of RCT data and that, in practice, patients who are not indicated for EGFR inhibitors (those with *KRAS* mutations) do not receive EGFR inhibitors but are given potentially beneficial alternative treatments, and patients who are indicated for EGFR inhibitor treatment (i.e., those without *KRAS* mutations) are treated appropriately.[28]

Although there is some limited evidence on the cost effectiveness of EGFR inhibitors, there are no published cost-effectiveness analyses of the impact of *KRAS* testing. A recent unpublished analysis indicates that pre-treatment *KRAS* mutation testing for all new cases of metastatic colorectal cancer would result in theoretical savings of hundreds of millions of dollars annually in the avoidance of ineffective treatment with cetuximab for patients with *KRAS* mutations. Further analyses based on empirical data may provide better estimates of cost savings and cost effectiveness of *KRAS* testing.

As some *wild-type KRAS* patients do not respond to EGFR inhibitors, further research is needed to improve identification of those patients (perhaps using a combination of genetic tests influencing the EGFR signaling pathway), in order to better direct their treatment selection. Also, *KRAS* mutation testing to date is generally limited to patients with metastatic colorectal cancer. There is growing interest in its use to inform treatments in patients with early stage disease, similar to the expansion of *HER-2/neu* testing for patients with breast cancer. Further research in expanded use of *KRAS* testing will be important in realizing potential further gains in quality of care, patient outcomes, and cost savings in the treatment of colorectal cancer.

References

1. Erbitux (cetuximab). Package insert. Imclone Systems Inc., 2007. http://packageinserts.bms.com/pi/pi_erbitux.pdf.)

2. Parkin DM, Bray F, Ferlay J, et al. Global cancer statistics. Washington, DC: American Cancer Society, 2005.

3. Jansman F, Postma MJ, Brouwers J. Cost considerations in the treatment of colorectal cancer. Pharmacoeconomics 2007;25(7):537-62.

4. Ramsey SD, Berry K, Etzioni R. Lifetime cancer-attributable coat of care for long-term survivors of colorectal cancer. Am J Gastroenterol 2002;97:440-5.

5. Avastin (bevacizumab). Package insert. Genetech, 2009. http://www.gene.com/gene/products/information/pdf/avastin-prescribing.pdf.)

6. Vectibix (panitumumab). Package insert. Amgen, 2007. http://www.amgen.com/pdfs/misc/vectibix_pi.pdf.)

7. Evans WE, McLeod HL. Pharmacogenomics--drug disposition, drug targets, and side effects. N Engl J Med 2003;348(6):538-49.

8. Piper MA, Aronson N, Ziegler KM, et al. Special report: pharmacogenomics of cancer--candidate genes. Report No. 22(5). Technology Evaluation Center, Blue Cross Blue Shield Association, 2007.

9. Conlin A, Smith G, Carey FA, et al. The prognostic significance of K-ras, p53, and APC mutations in colorectal carcinoma. Gut 2005;54:1283-6.

10. Linardou H, Dahabreh J, Kanaloupiti D, et al. Assessment of somatic *k-RAS* mutations as a mechanism associated with resistance to EGFR-targeted agents: a systematic review and meta-analysis of studies in advanced non-small-cell lung cancer and metastatic colorectal cancer. The Lancet Oncology 2008;9(10):962-72.

11. Sartore-Bianchi A, Martini M, Molinari F, et al. *PIK3CA* mutations in colorectal cancer are associated with clinical resistance to EGFR-targeted monoclonal antibodies. Cancer Res 2009;69(5):1851-7.

12. Moore MJ, Goldstein D, Hamm J, et al. Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. J Clin Oncol 2007;25:1960-6.

13. Massarelli E, Varella-Garcia M, Tang X, et al. KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. Clin Cancer Res 2007;13:2890-6.

14. Amado RG, Wolf M, Peeters M, et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. J Clin Oncol 2008;26:1626-34.

15. Bokemeyer I, Bondarenko JT, Hartmann FG, et al. KRAS status and efficacy of first-line treatment of patients with metastatic colorectal cancer (mCRC) with FOLFOX with or without cetuximab: the OPUS experience. J Clin Oncol 2008;26(Supplement):4000.

16. Karapetis CS, Khambata-Ford S, Jonker DJ, et al. K-ras mutations benefit from cetuximab in advance colorectal cancer. N Engl J Med 2008;359:1757-65.

17. Punt CJ, Tol J, Rodenburg CJ, et al. Randomized phase III study of capecitabine, oxaliplatin, and bevacizumab with or without cetuximab in advanced colorectal cancer (ACC), the CAIRO2 study of the Dutch Colorectal Cancer Group (DCCG). J Clin Oncol 2008;26(Supplement):abstr LBA4011.

18. Van Cutsem E, Lang I, D'haens G, et al. KRAS status and efficacy in the first-line treatment of patients with metastatic colorectal cancer (mCRC) treated with FOLFIRI with or without cetuximab: the CRYSTAL experience. J Clin Oncol 2008;26(Supplement):2.

19. Au HJ, Karapetis CS, O'Callaghan CJ, et al. Health-related quality of life in patients with advanced colorectal cancer treated with cetuximab: overall and KRAS-specific results of the NCIC CTG and AGITG CO.17 Trial. J Clin Oncol 2009;27(11):1822-8.

20. De Roock W, Piessevaux H, De Schutter J, et al. KRAS wild-type state predicts survival and is associated to early radiological response in metastatic colorectal cancer treated with cetuximab. Ann Oncol 2008;19:508-15.

21. Van Cutsem E, Kohne CH, Hitre R, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. N Engl J Med 2009;360(14):-1408.

22. CTC v2.0 and common terminology criteria for adverse events (CTCAE). National Cancer Institute, 2009. (Accessed June 22, 2009, at http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm.)

23. Realizing the promise of pharmacogenomics: opportunities and challenges. Draft report of the Secretary's Advisory Committee on Genetics, Health, and Society. Bethesda, MD: Secretary's Advisory Committee on Genetics, Health, and Society, 2007.

24. NCCN clinical practice guidelines in oncology: colon cancer. National Comprehensive Cancer Network, 2009. (Accessed June 22, 2009, at http://www.nccn.org/professionals/physician_gls/PDF/colon.pdf.)

25. Allegra CJ, Jessup JM, Somerfield MR, et al. American Society of Clinical Oncology provisional clinical opinion: testing for KRAS gene mutations in patients with metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy. American Society of Clinical Oncology, 2009. (Accessed July 20, 2009, at http://www.asco.org/ASCOv2/Practice+&+Guidelines/Guidelines/Provisional+Clinical+Opinion.)

26. Cetuximab (Erbitux) and panitumumab (Vectibix). Class labeling changes to anti-EGFR monoclonal antibodies, cetuximab (Eribitux) and panitumumab (Vectibix): KRAS mutations. Food and Drug Administration, 2009. (Accessed July 20, 2009, at http://www.fda.gov/aboutfda/centersoffices/CDER/ucm172905.htm.)

27. Brown HM, Aronson N, Ziegler KM, et al. KRAS mutations and epidermal growth factor receptor inhibitor therapy in metastatic colorectal cancer. Report No. 23(6). Technology Evaluation Center, Blue Cross Blue Shield, 2009.

28. Colorectal cancer (first line) - cetuximab: appraisal consultation document. London: National Institute for Health and Clinical Excellence, 2009. http://www.nice.org.uk/guidance/index.jsp?action=article&o=42073.

29. Merck Serono response to NICE ACD: Cetuximab in the treatment of metastatic colorectal cancer (mCRC). London: National Institute for Health and Clinical Excellence, 2008. http://www.nice.org.uk/nicemedia/pdf/ColorectalCancerCetuxamibACD2ManMerck.pdf.

30. Mittmann N, Au HJ, Tu D, et al. Prospective cost-effectiveness analysis of cetuximab in metastatic colorectal cancer: evaluation of National Cancer Institute of Canada Clinical Trials Group CO.17 Trial. J Natl Cancer Inst 2009;101(17):1182-92.

31. Shankaran V, Bentrem DJ, Mulcahy MF, et al. Economic implications of KRAS testing in metastatic colorectal cancer (mCRC). 2009. American Society of Clinical Oncology. 7-20-2009.

Case Study IV:

Screening for Cervical Cancer with HPV Genetic Test

Magnitude and Importance of Cervical Cancer

Cervical cancer is the second most commonly diagnosed and third leading cause of cancer mortality among women worldwide.[1] More than 555,000 new cases of cervical cancer are diagnosed annually across the globe, with more than 80% of cases occurring in developing countries.[2] The highest incidence rates are in Central and South America, the Caribbean, Sub-Sahara Africa, and Southern Asia. The disproportionate burden of cervical cancer in developing countries and elsewhere in medically underserved populations is mainly due to lack of access[1] to effective screening programs.[2, 4]

In the United States, about 10,800 cases of HPV-associated cervical cancer are diagnosed each year, of which approximately 3,700 cases result in death, according to the CDC, based on incidence data for 1998-2003.[5] The American Cancer Society (ACS) projects that approximately 11,270 new cases of invasive cervical cancer will be diagnosed and that there will be 4,070 deaths from the disease in 2009.[6] Incidence and mortality rates have steadily declined over the past several decades, from 3.7 deaths per 100,000 in 1986 to 2.4 deaths per 100,000 in 2005, as a result of emphasis on prevention and early detection through increased screening.[7, 8] The decline is largely attributable to the widespread acceptance use of cervical cytology screening tests, including conventional Papanicolaou ("Pap") tests and newer forms of cervical cytology.[2] Cervical cytology tests served as the gold standard and primary means of screening for cervical cancer since the 1940s.

Scientific advances in the 1980s and 1990s led to recognition that human papillomaviruses (HPVs) are the primary cause of cervical cancer and to two innovations: HPV assays for screening and the HPV vaccine.[9-11] Together, prevention using the HPV vaccine and detection using cervical cytology provide opportunities for greater and more efficient reductions in the incidence of cervical cancer and associated costs of care.

More than 150 different types of HPV-related viruses have been identified, of which about 30 are associated with cervical carcinomas.[12, 13] Low-risk types associated with other types of HPV infection generally clear spontaneously without intervention. Epidemiological studies on the global prevalence of HPV types indicate that about 70% of cervical cancer cases are related to infections with two high-risk HPV types[3]: HPV 16 and HPV 18.[14, 15] About 15% of cervical cancers are related to high-risk HPV types 31, 33, 35, 45, 52, and 58, and another 15% are less

---

[1] Factors that contribute to insufficient screening programs in developing countries include over-reliance on maternal and child health services for screening, as their target population is generally too young, opportunistic rather than organized screening, and low coverage of the target group.[3]

[2] The decline is most notable in squamous (epithelial) cell carcinomas, which can be detected with Pap tests; however, the tests are less effective in detecting adenocarcinomas (cancer originating in glandular tissue).

[3] HPV vaccines are designed to prevent infections caused by the two high-risk types, HPV 16 and HPV 18, as well as other common types that cause genital warts, HPV 6 and HPV 11.

common types with some variation among geographical regions.[14, 15] HPV-related cervical carcinomas develop as the result of long-term, persistent infection with high-risk types.[16]

Significant racial and ethnic disparities persist with regard to incidence and mortality rates of cervical cancer.[5] According to CDC analyses of data for 1998-2003, the cumulative incidence rate for all types of cervical cancers was higher in African American women (12.6%) compared to white women (8.4%) and Asian/Pacific Island (API) women (8.3%). Hispanic women also had increased incidence (14.2%) compared to non-Hispanics (8.4%). African American women tended to be diagnosed at an older median age (49 years vs. 47 years for white women) whereas Hispanic women tended to be diagnosed at a younger age (45 years vs. 48 years for non-Hispanics). For each race/ethnic group, most localized cervical carcinomas occurred in women under 50 years, particularly those ages 40-49 years. After 50 years, cervical cancers of all stages decreased for white women, but remained high for African Americans; for Hispanic and API women over 50 years, rates of localized cancer declined but regional and advanced stage cancers increased. These higher incidence rates for African Americans and Hispanics have been attributed to less screening or follow-up after abnormal test results.[17-19] these differences in screening rates may reflect differences in access to health care services more broadly due to distance to health care facility, language or other cultural barriers, and lack of health insurance.[5]

Laboratory Tests for Diagnosing and Treating Cervical Cancer

Currently, there are several types of laboratory tests for screening and diagnosis of cervical cancer, including **cervical cytology** and innovative **molecular HPV DNA tests**. These tests aim to detect potentially pre-cancerous changes, known as cervical intraepithelial neoplasia[4] (CIN) or cervical dysplasia, usually caused by sexually transmitted HPVs. The HPV test also can indicate whether atypical squamous cells of undetermined significance (ASCUS) collected for cytology testing have the potential to progress to cancer. CIN and ASCUS are important indicators of molecular-level changes in cervical cells that, if left untreated, can lead to cancer.

Traditional methods of cervical cytology testing first established in the 1940s involve the collection of cells from the cervix, affixing the specimen to a glass slide (via fixative spray) for microscopic evaluation to determine the presence of abnormalities. In the 1990s, innovations in cytology testing techniques led to the development of liquid-based thin-layer slide preparation testing method whereby the specimen is mixed into a vial of liquid preservative and then made into slide samples.[20] Thin-layer methods provide for cleaner and more uniform analysis, improving testing accuracy and efficiency over traditional methods.[21] In 2007, 99% of laboratories used thin-layer cervical cytology test methods for cervical cancer screenings.[20, 22] These tests remain an effective and widely used means of detecting certain cervical precancers and cancers.

Further innovations have been realized with the discovery of cervical cancer biomarkers and corresponding development of molecular testing techniques to screen for HPV DNA. Two

---

[4] Cervical intraepithelial neoplasia (CIN) refers to a method of classifying the spectrum of abnormality (dysplasia) in cervical lesions to help standardize treatment. CIN1 is mild dysplasia, CIN2 is moderate dysplasia, and CIN3 is severe dysplasia. The most recent classification is the Bethesda System, divides all cervical epithelial presursor lesions into 2 groups: Low-grade Squamous Intraepithelial Lesion (LSIL) and High-grade Squamous Intraepithelial Lesion (HSIL). LSIL corresponds to CIN1, and HSIL includes CIN2 and CIN3. More recently, CIN2 and CIN3 have been combined into CIN2/3.

predominant types of molecular tests are available to screen for high-risk HPV DNA: the **hybrid capture II assay** (HC2) (i.e., direct hybridization) and **PCR-based method** (i.e., amplification). HC2 can distinguish between groups of high-risk and low-risk genotypes, but it does not permit identification of individual genotypes.[23, 24] PCR methods amplify parts of the HPV DNA genome to support type-specific identification (including high-risk HPV 16 and 18), but may underestimate the prevalence of multiple genotypes when used for broad spectrum testing (high-risk vs. low-risk).[25]
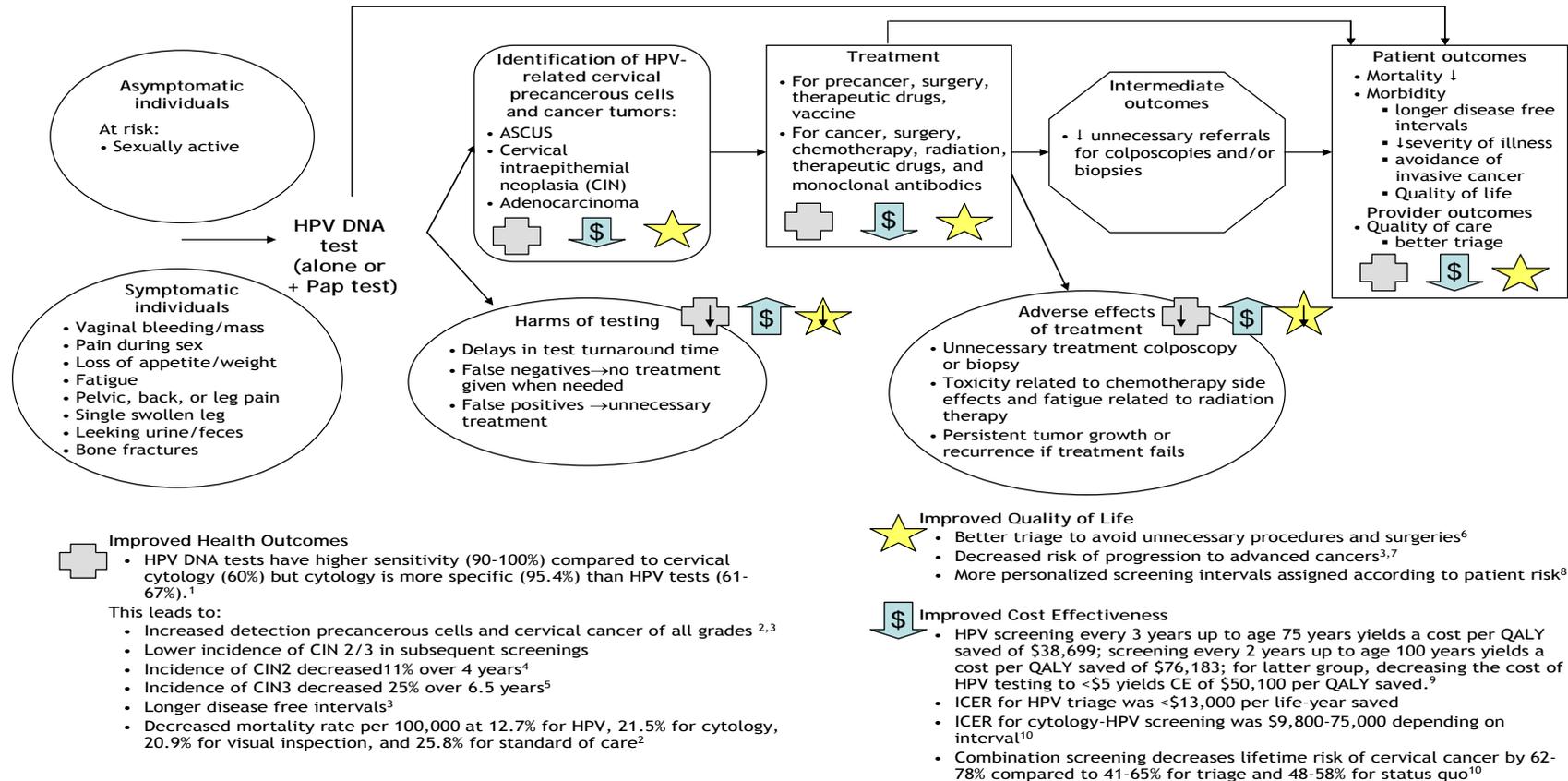
While both testing methods were developed in-house by laboratories (laboratory developed tests), one HC2 test also has received FDA approval.[26] In March 2009, FDA approved an innovative isothermal signal amplification technology.[27] Other emerging molecular-based technologies include the Luminex xMap to detect an expanded range of HPV; ThirdWave genotype assay and DNA chip probes for type-specific detection; viral load/real-time PCR to distinguish relevant infection; a rapid, simple assay to detect HPV16; and ProExC to identify markers of aberrant S-phase induction.[24]

Clinical Value of HPV Screening

This case study focuses on the clinical and economic value of HPV testing in cervical cancer screening programs as presented in the clinical analytic framework (Figure 1). HPV DNA testing can yield significant **clinical value** in several ways, including the following:

- Improved diagnostic accuracy and predictive value in identifying individuals with precancer ASCUS as well as cervical intraepithelial neoplasia (CIN)

- More effective triage of patients after cytology, decreasing effect of cytology false-negative test results

- Improved potential to prevent cervical cancer by identifying precancerous states

- Improved early detection and prediction of therapeutic outcome

- Decreased need for surgical procedures

- Improved patient outcomes, including decreased mortality and morbidity (e.g., surgical procedures, chemotherapy), and increased quality of life

- Decreased costs of cancer as a result of early intervention and cancer prevention cancer, and/or reduction in and severity of illness and complications

Figure 1. Clinical Analytic Framework for HPV Testing

Asymptomatic individuals

At risk:
- Sexually active

HPV DNA test (alone or + Pap test)

Symptomatic individuals
- Vaginal bleeding/mass
- Pain during sex
- Loss of appetite/weight
- Fatigue
- Pelvic, back, or leg pain
- Single swollen leg
- Leeking urine/feces
- Bone fractures

Identification of HPV-related cervical precancerous cells and cancer tumors:
- ASCUS
- Cervical intraepithemial neoplasia (CIN)
- Adenocarcinoma

Treatment
- For precancer, surgery, therapeutic drugs, vaccine
- For cancer, surgery, chemotherapy, radiation, therapeutic drugs, and monoclonal antibodies

Intermediate outcomes
- ↓ unnecessary referrals for colposcopies and/or biopsies

Patient outcomes
- Mortality ↓
- Morbidity
  - longer disease free intervals
  - ↓severity of illness
  - avoidance of invasive cancer
- Quality of life

Provider outcomes
- Quality of care
  - better triage

Harms of testing
- Delays in test turnaround time
- False negatives→no treatment given when needed
- False positives →unnecessary treatment

Adverse effects of treatment
- Unnecessary treatment colposcopy or biopsy
- Toxicity related to chemotherapy side effects and fatigue related to radiation therapy
- Persistent tumor growth or recurrence if treatment fails

Improved Health Outcomes
- HPV DNA tests have higher sensitivity (90-100%) compared to cervical cytology (60%) but cytology is more specific (95.4%) than HPV tests (61-67%).[1]

This leads to:
- Increased detection precancerous cells and cervical cancer of all grades [2,3]
- Lower incidence of CIN 2/3 in subsequent screenings
- Incidence of CIN2 decreased11% over 4 years[4]
- Incidence of CIN3 decreased 25% over 6.5 years[5]
- Longer disease free intervals[3]
- Decreased mortality rate per 100,000 at 12.7% for HPV, 21.5% for cytology, 20.9% for visual inspection, and 25.8% for standard of care[2]

Improved Quality of Life
- Better triage to avoid unnecessary procedures and surgeries[6]
- Decreased risk of progression to advanced cancers[3,7]
- More personalized screening intervals assigned according to patient risk[8]

Improved Cost Effectiveness
- HPV screening every 3 years up to age 75 years yields a cost per QALY saved of $38,699; screening every 2 years up to age 100 years yields a cost per QALY saved of $76,183; for latter group, decreasing the cost of HPV testing to <$5 yields CE of $50,100 per QALY saved.[9]
- ICER for HPV triage was <$13,000 per life-year saved
- ICER for cytology-HPV screening was $9,800-75,000 depending on interval[10]
- Combination screening decreases lifetime risk of cervical cancer by 62-78% compared to 41-65% for triage and 48-58% for status quo[10]

Sources:

1. Dillner J, Rebolj M, Birembaut P, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. Br Med J 2008;337:1-8.
2. Mayrand MH, Duarte-Franco E, Rodrigues I, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. N Engl J Med 2007;357(16):1579-88.
3. Cuzick J, Arbyn M, Sankaranarayanan R, et al. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. Vaccine 2008;26S:K29-K41.
4. Naucler P, Ryd W, Tomberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. N Engl J Med 2007;357:1589-97.
5. Bulkmans NW, Berkhof J, Rozendaal L, et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomized controlled implementation trial. Lancet 2007;370:1764-72.
6. Gravitt P, Coutlee F, Iftner T, et al. New technologies in cervical cancer screening. Vaccine 2008;26((Supplement 10)):K42-K52.
7. Dehn D, Torkko KC, Shroyer KR. Human papillomavirus testing and molecular markers of cervical dysplasia and carcinoma. Cancer (Cancer Cytopathology) 2007;111(1):1-14.
8. Petry KU, Menton S, Menton M, et al. Inclusion of HPV testing in routine cervical cancer screening for women above 29 years in Germany: results for 8,466 patients. Br J Cancer 2003;88:1570-7.
9. Mandelblatt JS, Lawrence WF, Womack SM, et al. Benefits and costs of using HPV testing to screen for cervical cancer. JAMA 2002;287(18):2372-81.
10. Kim JJ, Wright TC, Goldie SJ. Cost-effectiveness of human papillomavirus DNA testing in the United Kingdom, the Netherlands, France, and Italy. J Natl Cancer Inst 2005;97(12):888-95.

Detection

There is substantial evidence demonstrating that the sensitivity (ability to detect true positives) of HPV testing (both HC2 and PCR) is very high, but the specificity (ability to detect true negatives) is lower compared to cytology. Pooled cumulative estimates of the diagnostic accuracy are presented in Table 1. Diagnostic accuracy for all cervical cancer screening tests can vary slightly according to cancer stage. The high specificity of cytology demonstrates advantages for confirming negative results while the high sensitivity of HPV tests demonstrates advantages in identifying individuals with ASCUS as well as cervical intraepithelial neoplasia (CIN) grades 2/3 or higher.[24, 28]

Some studies have reported increased sensitivity from the liquid-based Pap tests compared to conventional Pap tests, while others found no statistically significant differences.[29-32] In addition, use of liquid-based Pap tests does not change ASCUS detection rates for follow-up reflex HPV testing.[33]

Table 1. Sensitivities, specificities, and positive/negative predictive values (PPV, NPV) reported in the literature for cervical cancer screening/diagnostic tests.

Test	Sensitivity	Specificity	NPV	PPV
Cytology	• 59.9% • 61-66% (thin-layer)	• 95.4% • 82-91% (thin-layer)	• 99%	• 21.1%
HC2	• ≥96% • 93.6% for CIN3 or cancer • 70% for cervical lesions	• 61% • 41.2% for CIN3 or cancer • 56% for cervical lesions	• 99% • 98.2% for CIN3 or cancer • 95% for cervical lesions	• 23% for CIN2+ • 16.1-17.2% for CIN3 or cancer • 19% for cervical lesions
PCR amplification	• 92% for CIN2+ using Gp5+/6+ • 89.9-100% for CIN2+ and 89.3% for CIN3 using Amplicor	• 66.7% for CIN2+ using Gp5+/6+ • 35.7-49.2% for CIN2+ and 48.5% for CIN3 using Amplicor	• 85.7% for CIN2+ using Gp5+/6+ • 96.7% for CIN2+ and 97.4% for CIN3 using Amplicor	• 81.3% for CIN2+ using Gp5+/6+ • 33.7% for CIN2+ and 17.3% for CIN3 using Amplicor

Sources:

Coste J, Cochand-Priollet B, de Cremoux P, et al. Cross sectional study of conventional cervical smear, monolayer cytology, and human papillomavirus DNA testing for cervical cancer screening. BMJ 2003;326 (7392):733.

Dehn D, Torkko KC, Shroyer KR. Human papillomavirus testing and molecular markers of cervical dysplasia and carcinoma. Cancer 2007;111(1):1-14.

Dillner J, Rebolj M, Birembaut P, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. Br Med J 2008;337:1-8.

Kulasingam SL, Hughes JP, Kiviat NB, et al. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: comparison of sensitivity, specificity, and frequency of referral. JAMA 2002;288 (14): 1749-57.

In an observational study of HPV testing in 167 women, both types of HPV tests (HC2 and PCR) yielded comparable results with a good degree of concordance (83.3%), indicating that both are suitable for routine use.[34] Within the 17% of discordant results, there are some modest differences among the tests. One study found that, for PCR+/HC2- results, the higher sensitivity of the PCR method identified 12% more positive results; of these, PCR correctly

identified HPV in 65%, but incorrectly identified the other 35% as positive (i.e., false-positive results).[34] Conversely, for HC2+/PCR- results, the PCR method will correctly verified that no HPV was present. Another important measure is the rate of false-negative results. PCR correctly identified all patients with CIN3, but HC2 failed to detect it in two of seven patients

Several randomized controlled trials (RCTs) compared HPV- and cytology-based methods as primary cervical cancer screening tests.[35] Data from the trials confirm that HPV-based screening results in detection of more high-grade CIN lesions but reduced specificity compared to cytology-based screening. These trials also found that the increased sensitivity for CIN3+ (of the HPV tests) corresponded correctly to a lower incidence of CIN3+ in future screening.[36, 37]

Thus, all three types of cervical cancer screening tests have implications for clinical practice. At present, HPV testing is generally used in combination with cervical cytology to screen for HPV infection as well as triage borderline Pap test results. HPV tests in combination with Pap tests are 96% to 100% sensitive for detection of CIN 2/3 and cancer.[38] Because of its lower false-positive rate, HC2 is often used with cytology for routine screening. However, the higher specificity of PCR for detecting CIN3 can provide a better triage for borderline results.[34, 39, 40]

Clinical Outcomes

Along with improved detection, HPV testing for cervical cancer yields several other benefits that improve clinical care. These benefits include a better means of triaging patients prior to referral for colposcopy or biopsy, longer disease-free intervals and other improved long-term health outcomes, and avoidance of invasive cancers and decreased cervical cancer mortality. The following section summarizes key findings in the literature related to these important outcomes.

Triage and referral for additional procedures

Screening for ASCUS via Pap testing involves a significant degree of uncertainty and can vary significantly among laboratories. The sensitivity of HPV DNA molecular testing provides an objective means to triage patients to confirm positive or negative results.[41] For example, triage is important for women who are referred for colposcopy because of abnormal cytology tests but who do not have any visible lesion or for whom a biopsy yields negative histology. For these women, a negative HPV test at or near the time of colposcopy provides additional reassurance that there is unlikely to be any undetectable disease. A positive HPV result, especially for high-risk types 16 and 18, provides assurances of continued risk and need for short-term repeat testing and closer patient follow-up.[42] The value of HPV DNA testing in triage is based in part on avoiding unnecessary procedures that would have occurred in the absence of this test. For example, if all women would otherwise have been referred for colposcopy following abnormal Pap tests, then HPV testing can help to reduce rates of unnecessary biopsy referrals.[43]

Longer disease-free survival and improved long-term outcomes

Several studies have shown that negative HPV results alone or in combination with negative cytology signifies predicts longer disease-free interval against CIN2+ than negative cytology results alone.[43] In particular, long-term follow-up of two recent large RCTs in Sweden and the Netherlands demonstrated that the higher detection rate for CIN2+ with HPV, used as part of screening, led to lower rates of CIN3+ at subsequent screening and indicated that HPV DNA tests are highly sensitive to prevalent cases. In the Swedish study of 12,527 women ages 32-38,

the addition of HPV testing at the initial screen increased detection rate for CIN2+ by 51% compared to Pap test only and, in subsequent screenings over a 4-year (mean) follow-up period, the proportion of women found to have grade 2 or 3 lesions was 42% less and the proportion of grade 3 lesions or cancer was 47% less.[37] Similar results were found in the Dutch study of 8,575 women, where the addition of HPV testing increased initial detection rate of CIN3+ by 70% compared to Pap test only, and the number of grade 3 lesions was 55% lower over the 6.5 year (mean) follow-up period.[36] The investigators concluded that HPV testing leads to earlier detection of CIN3+ lesions, effectively lowering incidence CIN3+ at 6.5 years and permitting an extension of the screening interval.

Many observational studies have demonstrated that HPV DNA testing is a good predictor of risk. Protection over an extended period also was demonstrated in an earlier 2003 German study, the first large-scale evaluation of HPV testing as an adjunct to routine cytology screening.[44] Combination screening enabled population risk stratification and development of risk-adapted screening programs with increased intervals for those at lower risk. Study results demonstrated that HPV+/Pap+ test results identified the highest proportion of high-grade cervical neoplasia (19/54). Similarly, women with HPV-/Pap- test results are at significantly reduced risk of developing high-grade CIN or cancer, suggesting that they should be protected for an extended period of time between screenings.

A recent (2008) multinational study pooled data on 24,295 women to determine the effects of HPV testing compared to cytology at a six-year follow up.[35] Analyses determined that the PPV for future CIN3+ was highest at 34% among women with abnormal cytology and positive HPV test at baseline. Women with normal cytology but positive HPV had continuously increasing cumulative incidence rate of CIN3+, eventually reaching 10% after six years, whereas women with abnormal cytology and negative HPV had an incidence rate of 2.7%. At six-year follow-up, the rate of CIN3+ was significantly lower among women negative for HPV (0.12-0.45%) than women negative for cytology results (0.53-1.34%). Results were similar for incidence of CIN grade 2 or worse.[35]

In addition, HPV testing is a sound tool for lowering risk of cervical cancer. Another comparative observational study of 2,982 women found that the proportion of women with CIN2+ within 1, 5, and 9 years after negative HPV test (0.19%, 0.42% and 1.88% respectively) was half as high as that for women with negative cytology (0.33%, 0.83% and 2.20%).[45] HPV testing offered excellent protection from CIN2+ for at least 6 years after a negative test, whereas the protection from cytology began to wane after about 3 years.

These data on decreased incidence of advanced-stage cancers, along with data from numerous other clinical trials and observational studies noted earlier, have implications for the ability of HPV testing to contribute to decreased mortality. The potential for greater impact on mortality was demonstrated in a recent computer-based simulation. According to this simulation, adding HPV testing to Pap screening produces other substantial clinical benefits, such as avoidance of 225 invasive cancers per 100,000 women and decreases cervical cancer mortality by 59%.[46] (See further discussion below on economic value.)

Efficient primary screening test

HPV DNA testing is increasingly recognized for its potential as a primary screening test for cervical cancer alone and as the initial test followed by cytology for triage. In early 2009, the results of a nine-year RCT conducted in rural India reported on the effect of a single round of screening by HPV DNA, cytology testing, or visual inspection of the cervix with ascetic acid (VIA) on the incidence of cervical cancer and mortality compared to standard care.[4] The RCT evaluated 131,746 women ages 30-59 years who were randomly assigned to a screening group or the control group (which received the standard of care). Women who had positive results were referred for colposcopy and directed biopsies and those with precancerous lesions were given appropriate treatment. Study results demonstrated that incidence of CIN2+ cancer per 100,000 populations and death were significantly lower in the HPV testing group compared to cytology and VIA testing. Investigators concluded that a single round of HPV testing was associated with a significant reduction in the number of CIN2+ cancers and cervical cancer-related deaths.

More specifically, the incidence of CIN2+ cancer at the eight-year follow up was 14.5% for HPV, 23.3% for cytology, 32.2% for VIA, and 33.1% for the control group.[4] Mortality rate per 100,000 followed a similar pattern at 12.7% for HPV testing, 21.5% for cytology, 20.9% for VIA, and 25.8% for the control group. Throughout the eight-year follow up period, the gap between incidence of advanced cancer and mortality for HPV testing and that for the control group widened significantly. Investigators concluded that a single round of HPV testing was associated with a significant decline in the rate of advanced cervical cancers and associated deaths, as compared to the unscreened group. The mortality rate did not change for the cytology or VIA groups. The investigators also found that HPV testing was the most objective and reproducible of all cervical cancer screening tests and was less demanding in terms of staff training and quality assurance. Given the results of the study, the investigators stated that, in developing countries, HPV screening could begin at 30 years and be performed once every 10 years. However, the extensive interval would not be accepted in developed countries where routine screenings are performed annually or at 2-3 year intervals, especially since the HPV test may indicate a reduced risk, not zero risk.[46, 47] Several studies (including those cited above) have examined the influence of screening intervals, suggesting that longer intervals are associated with similar or increased detection of high-grade lesions.[48, 49]

Available evidence suggests that the use of HPV testing as the sole primary screening modality provides certain benefits:

(1) HPV DNA detection assays provide an automated, objective, and very sensitive test allowing for better quality control and reduces the basis for medical legal claims;

(2) Cytology can be reserved for the 5-15% of women who are HPV positive, improving the quality and efficiency of cytology screening;

(3) HPV screening avoids unnecessary triage of HPV-negative ASCUS/LSIL; and

(4) If validated with well-designed clinical trials, longer screening intervals may be safe, which would improve the cost and convenience of screening.[36, 37, 50]

However, because non-cancerous HPV infections are common, especially in young women, which will clear spontaneously, a drawback of HPV DNA testing as the primary means of screening for cervical cancer may be over-response to the detection of HPV in the form of

unnecessary colposcopies, psychological distress, and unnecessary treatment.[43] In addition, for low resource areas, there is concern about the costs associated with screening regimens involving combination screening versus Pap or HPV testing alone. The Indian study mentioned above cited the cost of the HPV test as a drawback, particularly for low resource settings. The HPV test is more expensive ($20-30 per test) and time consuming than other screening methods, and requires a sophisticated laboratory infrastructure. To address this need, in part, researchers are developing simple, affordable, and accurate HPV tests that provide results within three hours.

A Randomized Trial in Screening to Improve Cytology (ARTISTIC) Trial evaluated the effectiveness of HPV testing in more than 25,000 women undergoing routine cervical cancer screening. The trial included a cross-sectional study of HPV testing at the recruitment phase. The study found that HPV prevalence increased with level of cytological abnormality, more so at older ages. About 87% of women under 30 years with mild cytologic abnormalities were HPV positive; this proportion fell to 58% in women 30-49 years, and 28% for those 50-64 years. However, for those aged 20-29 years, the number who were high-risk HPV positive was 52% greater than the number with abnormal cytology, suggesting a need for a second test prior to colposcopy to avoid unnecessary procedures. Because of the high prevalence of HPV infection in women under 30 years, the investigators questioned the utility HPV as a primary screening tool for this age group.

Despite the available body of evidence of international trials and reviews that affirm certain aspects of the value of HPV testing, it has received only limited approval as an adjunct to Pap cytologic screening, i.e., only by certain authoritative groups in the US. Refer to CDC's complete summary of the cervical cancer screening guidelines endorsed by these groups which can be found at http://www.cdc.gov/std/hpv/ScreeningTables.pdf. Generally, most groups recommend beginning HPV screening as an adjunct to cervical cytology after onset of sexual activity or age 21. ACS and the American College of Obstetricians and Gynecologists (ACOG) recommend HPV testing every 2-3 years for women ≥ 30 years with 3 negative cytology tests whereas U.S. Preventive Services Task Force (USPSTF) recommends HPV testing every 3 years without conditions. Cessation of screening varies from >65 years (USPSTF) to ≥ 70 years with 3 consecutive negative tests (ACS).[51, 52]

In 2006, the American Society for Colposcopy and Cervical Pathology (ASCCP) together with other professional medical societies developed guidelines for managing women with ASCUS, CIM, and adenocarcinoma. The consensus guideline states that HPV DNA testing for high-risk types, repeat cervical cytology testing, or colposcopy are all acceptable methods for managing women with ASCUS.[53] Use of HPV DNA tests for reflex testing of ASCUS Pap test interpretations in women ≥ 30 years is a cost-effective method of confirming results and eliminates the need for patients to return to the physician's office. However, stratifying by HPV genotype (beyond detection) was not considered clinical beneficial by ASCCP since only 50% of CIN2+ lesions are associated with HPV 16 or 18 and the risk of CIN2+ among women with nononcogenic HPV 16/18 remained high enough to warrant colposcopy.[54]

At the time of its last (2003) review of cervical cancer screening, including HPV DNA testing, USPSTF concluded that the evidence was insufficient to recommend for or against the routine use of new technologies as a primary screening tool (i.e., HPV testing, liquid-based cytology, computerized rescreening, and algorithm based screening) or to determine whether new technologies are more effective than conventional Pap smear screening in reducing incidence of

or mortality from invasive cervical cancer.[51] Currently, USPSTF is revisiting the topic of cervical cancer screening and, as such, is expected to consider new evidence from several RCTs published since 2002 confirming the value of adding HPV testing to cervical cytology.[55]

Economic Value of HPV Screening

Few analyses have been conducted on the cost effectiveness of HPV testing. Most of these have been simulations of alternative screening strategies. An extensive simulation published in 2002 examined 18 cervical cancer screening strategies of Pap testing plus HPV testing, Pap testing alone, or HPV testing alone every two or three years among hypothetical cohorts of women beginning at age 20 and continuing until 65 years, 75 years, or death. Among the progression of most efficient scenarios (excluding those that were less cost effective or were both more costly and saved fewer lives than an alternative scenario, i.e., the "efficiency frontier"), the maximal savings in life years was achieved by using combined Pap-HPV screening versus Pap testing alone every two years without an upper age limit (up to age 100), for an incremental cost-effectiveness ratio (ICER) of $76,183 per QALY. Discontinuing screening with HPV plus Pap testing at age 75 still captured 97.8% of the benefits and lowered the ratio to $70,347 per QALY. Further along the efficiency frontier were Pap testing only every two years to 100 years ($56,440 per QALY), Pap testing only every two years to 75 years ($29,781 per QALY), and Pap only every three years to 75 years ($11,830 per QALY). Among the three-year testing scenarios, the Pap plus HPV screening up to age 75 years yielded a ratio of $38,699 per QALY. HPV screening alone was equally effective as Pap testing alone at any given screening interval or age of screening cessation but was more costly in each instance. According to sensitivity analyses, if the cost of HPV testing would drop to $5 per test, it would be more cost effective than Pap testing. The investigators noted that a combination of low HPV test cost, targeting to high-prevalence groups, and/or improved sensitivity would favor HPV as a primary screening strategy. They concluded that screening with HPV plus Pap testing every two years appears to save additional years of life at reasonable costs compared with Pap testing alone.[46]

One comparative analysis examined the cost effectiveness of incorporating HPV testing into existing cervical cancer screening programs in four European countries: UK, Netherlands, France, Italy.[56] The study involved a computer-based model[5] of country-specific data on age-specific cervical cancer risk and a comparison of each country's current screening policy[6] with two new strategies: (1) cytology throughout a woman's lifetime with use of HPV testing as a triage strategy for equivocal cytology results; and (2) cytology until age 30 years and HPV testing in combination[7] with cytology in women more than 30 years. Outcomes included reduction in lifetime cervical cancer risk, increase in life expectancy, lifetime costs, and ICER, expressed as cost per life year saved. Cost data were estimated from published literature and included direct medical costs (e.g., cost of screening test, treatment, staff time, and office visits) and patient time costs.

---

[5] The model input parameters for the natural history for cervical carcinogenesis were based on population studies primarily in the United States.

[6] Each country's status quo policy assumed conventional cytology at screening ages, intervals, and coverage rates per the respective country's national health insurance program.

[7] Outcomes were calculated using three- and five-year HPV screening intervals.

Results demonstrated that both HPV testing strategies were more cost effective than each country's status quo strategy.[56] Adjusted to 2008 dollars, the ICERs for HPV triage were less than $13,000 per life year saved and that for the combination strategy ranged from $9,800 to $75,000 per life year saved, depending on the screening interval. HPV testing also contributed to significant reductions in lifetime risk of cancer given typical conditions and rates for repeat testing, referral for colposcopy, and biopsy. Specifically, cancer risk was reduced by 62-78% with combination screening compared to 41-65% for triage screening and 48-58% for the status quo. The most cost-effective strategy across all four countries was combination screening with HPV testing at three-year intervals. The investigators concluded that expansion of current cervical cancer screening programs with the addition of HPV testing at regular intervals has the potential to improve patient health at reasonable cost.

Another modeling study was conducted in 2005 on the cost effectiveness of cervical cancer screening in five developing countries—India, Kenya, Peru, South Africa, and Thailand.[57] Synthesis of data from the literature simulated the natural history of HPV-induced cervical cancer. Investigators calculated ICERs to assess alternative screening strategies that were differentiated according to targeted age groups, number of clinical visits, frequency of screening, and type of screening test, i.e., visual inspection of the cervix, conventional cytology test (Pap test), HPV test, and combination visual inspection and HPV testing. Cost data included direct medical costs, patient time costs, transportation costs, and program costs. Country-specific reduction in lifetime risk of cancer ranged 25-31% with a single screening at 35 years (1- and 2-visit visual inspection); 30-36% reduction with 1- and 2-visit HPV testing; and 18-22% with 2- and 3-visit cytology screening. Overall, two screenings (at 35 and 40 years) reduced lifetime risk by an additional 40%; the addition of a third screening at 45 years yielded further reductions by 55%.[57] Lifetime costs and cost-effectiveness ratios associated with different screening strategies varied among countries, owing to differences in costs of labor, tradable goods, time, and transportation.

Conclusions

Recognition that about 30 types HPVs are the primary cause of cervical cancer has led to important innovations through the development of HPV assays for screening and the HPV vaccine prevention. These innovations have brought new opportunities to improve screening and prevention of cervical cancer. Diagnostic accuracy for all cervical cancer screening tests, including cervical cytology, can vary slightly according to cancer stage. The high specificity of cytology demonstrates advantages for confirming negative results while the high sensitivity of HPV DNA tests demonstrates advantages in identifying individuals with ASCUS as well as cervical intraepithelial neoplasia (CIN) grades 2/3 or higher. Substantial evidence, based particularly on the Dillner (2008), Nauclear (2007), Bulkmans (2007), and Sankaranarayanan (2009) RCTs, demonstrates that, in combination, Pap and HPV testing provide opportunities for greater and more efficient reductions in the incidence of cervical cancer and associated costs of care.

HPV DNA testing yields several benefits to clinical care, including a better means of triaging patients prior to referral for colposcopy or biopsy and predicting longer disease-free intervals, as well as improved long-term health outcomes such as avoidance of invasive cancers and decreased cervical cancer mortality. Some studies also have illustrated the value of using HPV testing as a primary cervical cancer screening tool for women <30 years in terms of decreased

incidence of invasive cancer and mortality. In addition, studies of HPV testing in several countries have demonstrated high levels of cost effectiveness. The most cost-effective strategy is HPV-Pap combination screening at three-year intervals up to age 75 years.

Significant strides have been made in establishing a solid evidence base on the benefits of adding HPV DNA testing to cervical cancer screening programs. Yet, further study is needed to more clearly establish a direct link between HPV DNA testing and decreased incidence of cervical cancer and mortality. To support the recent scientific and forthcoming clinical advance in prevention, detection, and treatment of cervical cancer, policymakers should ensure the appropriate level of health system support for patient access to these important tests.

References

1. Initiative for vaccine research (IVR): human papillomavirus. World Health Organization, 2006. (Accessed July 14, 2009, at http://www.who.int/vaccine_research/diseases/hpv/en/.)

2. Global cancer facts and figures, 2007. Atlanta: American Cancer Society, 2007.

3. Cervical cancer screening in developing countries. Report of a WHO consultation. Geneva: World Health Organization, 2002. http://whqlibdoc.who.int/publications/2002/9241545720.pdf.

4. Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. N Engl J Med 2009;360(14):1385-94.

5. Watson M, Saraiya M, Benard V, et al. Burden of cervical cancer in the United States, 1998-2003. Cancer (Supplement) 2008;113(10):2855-64.

6. Cancer facts and figures 2009. Atlanta: American Cancer Society, 2009. http://www.cancer.org/downloads/STT/500809web.pdf.

7. Lesson 4: organizing epidemiologic data. In: Principles of epidemiology in public health practice. Centers for Disease Control and Prevention, 2009.

8. Cervical cancer deaths per 100,000 women, 2005. Kaiser Family Foundation, 2005. http://www.statehealthfacts.org/comparemaptable.jsp?ind=475&cat=10.

9. zur Hausen H. Papillomavirus infections--a major cause of human cancers. Biochim Biophys Acta 1996;1288:F55-F78.

10. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999;189(1):12-9.

11. Lowy DR, Solomon D, Hildesheim A, et al. Human papillomavirus infection and the primary and secondary prevention of cervical cancer. Cancer (Supplement) 2008;113:1980-93.

12. de Villiers EM, Fauquet C, Broker TR, et al. Classification of papillomaviruses. Virology 2004;324:17-27.

13. Munoz N, Bosch FX, Castellsague M, et al. Against which human papillomavirus types shall we vaccinate and screen? the international perspective. Int J Cancer 2004;111:278-85.

14. Clifford GM, Smith JS, Plummer M, et al. Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis. Br J Cancer 2003;88:63-73.

15. Quint W, Pagliusi SR, Lelie N, et al. Results of the first World health organization international collaborative study of detection of human papillomavirus DNA. J Clin Microbiol 2006;44(2):571-9.

16. Schiffman M, Herrero R, Desalle R, et al. The carcinogenicity of human papillomavirus types reflects viral evolution. Virology 2005;337:76-84.

17. Swan J, Breen N, Coates RJ, et al. Progress in cancer screening practices in the United States: results from the 2000 National Health Interview Survey. Cancer 2003;97:1528-40.

18. Benard VB, Lawson HW, Eheman CR, et al. Adherence to guidelines for the follow-up of low-grade cytologic abnormalities among medical underserved women. Obstetrics and Gynecology 2005;105:1323-8.

19. Eggleston KS, Coker AL, Das IP, et al. Understanding barriers for adherence to follow-up care for abnormal Pap tests. J Womens Health 2007;16:311-30.

20. Jhala D, Eltoum I. Barriers to adoption of recent technology in cervical screening. Cytojournal 2007;4:16.

21. Cramer B. New Pap smear provides more accurate test results. Nashville, TN: Vanderbilt Medical Center, 2000. http://www.mc.vanderbilt.edu/reporter/index.html?ID=1053.

22. Terry M. Lab industry strategic outlook: market trends and analysis 2007. New York, NY: Washington G-2 Reports, 2007.

23. Perrons C, Jelley R, Kleter B, Quint W, Brink N. Detection of persistent high risk human papillomavirus infections with hybrid capture II and SPF10/LiPA. J Clin Virol 2005;32(4):278-85.

24. Dehn D, Torkko KC, Shroyer KR. Human papillomavirus testing and molecular markers of cervical dysplasia and carcinoma. Cancer (Cancer Cytopathology) 2007;111(1):1-14.

25. van Doorn LJ, Molijn A, Kleter B, Quint W, Colau B. Highly effective detection of human papillomavirus 16 and 18 DNA by a testing algorithm combining broad-spectrum and type-specific PCR. J Clin Microbiol 2006;44(9):3292-8.

26. FDA premarket approval database. Food and Drug Administration, 2009. (Accessed July 15, 2009, at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMA/PMASimpleSearch.cfm.)

27. FDA premarket approval database: Cervista HPV 16/18. Food and Drug Administration, 2009. (Accessed July 15, 2009, at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMA/PMASimpleSearch.cfm?db=PMA&id=13843#AOStatement.)

28. Koliopoulos G, Arbyn M, Martin-Hirsch P, et al. Diagnostic accuracy of human papillovavirus testing in primary cervical screening: a systematic review and meta-analysis of non-randomized studies. Gynecol Oncol 2007;104(1):232-46.

29. Ronco G, Cuzick J, Pierotti P, et al. Accuracy of liquid based versus conventional cytology: overall results of new technologies for cervical cancer screening randomised controlled trial. Br Med J 2007;335(7609):1-2.

30. Coste J, Cochand-Priollet B, de Cremoux P, et al. Cross sectional study of conventional cervical smear, monolayer cytology, and human papillomavirus DNA testing for cervical cancer screening. Br Med J 2003;326(7392):733.

31. Bandyopadhyay S, Austin RM, Dabbs D, Zhao C. Adjunctive human papillomavirus DNA testing is a useful option in some clinical settings for disease risk assessment and triage of females with ASC-H Papanicolaou test results. Archives of Pediatrics and Laboratory Medicine 2008;132(12):1874-81.

32. Arbyn M, Bergeron C, Klinkhamer P, et al. Liquid compared with conventional cervical cytology: a systematic review and meta-analysis. Obstetrics and Gynecology 2008;111(1):167-77.

33. Thrall MJ, Russell DK, Bonfiglio TA, Hoda RS. Use of the ThinPrep(R) Imaging System does not alter the frequency of interpreting Papanicolaou tests as atypical squamous cells of undetermined significance. CytoJournal 2008;5(10):1-6.

34. Sandri MT, Lentati P, Benini E, et al. Comparison of the Digene HC2 assay and the Roche AMPLICOR humap papillomavirus (HPV) test for detection fo high-risk HPV genotypes in cervical samples. Journal of Microbiology 2006;44:2141-6.

35. Dillner J, Rebolj M, Birembaut P, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. Br Med J 2008;337:1-8.

36. Bulkmans NW, Berkhof J, Rozendaal L, et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomized controlled implementation trial. Lancet 2007;370:1764-72.

37. Naucler P, Ryd W, Tomberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. N Engl J Med 2007;357:1589-97.

38. Cox JT. Human papillomavirus testing in primary cervical screening and abnormal Papanicolaou management. Obstet Gynecol Surv 2006;61((6 Supplement 1)):S15-S25.

39. Castle PE, Solomon M, Schiffman M, et al. Human papillomavirus type 16 infections and 2-year absolute risk of cervical precancer in women with equivocal or mild cutologic abnormalities. J Natl Cancer Inst 2005;97:1066-71.

40. Kahn MJ, Castle PE, Lorincz AT, et al. The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. J Natl Cancer Inst 2005;97:1072-9.

41. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. JAMA 2002;287:2114-9.

42. Gravitt P, Coutlee F, Iftner T, et al. New technologies in cervical cancer screening. Vaccine 2008;26((Supplement 10)):K42-K52.

43. Cuzick J, Arbyn M, Sankaranarayanan R, et al. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. Vaccine 2008;26S:K29-K41.

44. Petry KU, Menton S, Menton M, et al. Inclusion of HPV testing in routine cervical cancer screening for women above 29 years in Germany: results for 8,466 patients. Br J Cancer 2003;88:1570-7.

45. Cuzick J, Szarewski A, Mesher D, et al. Long-term follow-up of cervical abnormalities among women screened by HPV testing and cytology: Results from the Hammersmith study. Int J Cancer 2008;122(10):2294-300.

46. Mandelblatt JS, Lawrence WF, Womack SM, et al. Benefits and costs of using HPV testing to screen for cervical cancer. JAMA 2002;287(18):2372-81.

47. McNeil DG. DNA test outperforms Pap smear. *New York Times*. 2009.

48. Sawaya GF, Kerlikowske K, Lee NC, et al. Frequency of cervical smear abnormalities within 3 years of normal cytology. Obstetrics and Gynecology 2000;96:219-23.

49. Sigurdsson K. Trends in cervical intra-epithelial neoplasia in Iceland through 1995: evaluation of targted age groups and screening intervals. Acta Obstet Gynecol Scand;78:486-92.

50. Cuzick J, Szarewski A, Cubie H, et al. Management of women who test positive for high-risk types of human papillomavirus: the HART study. Lancet 2003;362(9399):1871-6.

51. Screening for cervical cancer. Recommendations and rationale. U.S.Preventive Services Task Force, 2003. (Accessed August 20, 2009, at http://www.ahrq.gov/clinic/3rduspstf/cervcan/cervcanrr.pdf.)

52. CDC Summary Tables of Cx Ca Screening Guidelines and New Technologies. American Society for Colposcopy and Cervical Pathology, 2009. (Accessed August 10, 2009, at http://www.cdc.gov/std/hpv/ScreeningTables.pdf.)

53. Wright TC, Massad LS, Dunton CJ, et al. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. Am J Obstet Gynecol 2007:346-55.

54. Clinical update: HPV genotyping. American Society for Colposcopy and Cervical Pathology, 2009. (Accessed August 10, 2009, at http://www.asccp.org/pdfs/consensus/clinical_update_20090408.pdf.)

55. U.S. Preventive Services Task Force. Topics in progress. Agency for Healthcare Research and Quality, 2009. (Accessed September 22, 2009, at http://www.ahrq.gov/clinic/uspstf/topicsprog.htm.)

56. Kim JJ, Wright TC, Goldie SJ. Cost-effectiveness of human papillomavirus DNA testing in the United Kingdom, the Netherlands, France, and Italy. J Natl Cancer Inst 2005;97(12):888-95.

57. Goldie SJ, Gaffikin L, Goldhaber-Fiebert JD, et al. Cost-effectiveness of cervical cancer screening in five developing countries. N Engl J Med 2005;353(20):2158-68.