

2006 Cardiovascular Biomarkers And Surrogate Endpoint Symposium

September 18-20, 2006

Bethesda North Marriott Hotel & Conference Center
Bethesda, Maryland

Integrating Biomarkers into Practice: The Need for Data on Clinical Outcomes

Gurvaneet Randhawa, M.D.

Center for Outcomes and Evidence

Agency for Healthcare Research and Quality

Silver Spring, MD

The mission of the US Agency for Healthcare Research and Quality (AHRQ) is to improve the quality, safety, efficiency and effectiveness of healthcare in the US, which is a fairly ambitious mission for a small agency. One way AHRQ accomplishes this mission is by improving the evidence-base for decision-making by clarifying the evidence of benefits and harms of clinical interventions, broadly defined as diagnostics and therapeutics. This is accomplished through several programs and initiatives, one example is a program called Effective Health Care which is detailed at www.effectivehealthcare.ahrq.gov.

The U.S. Preventive Services Task Force (USPSTF) is independent of, but supported by AHRQ. The USPSTF is a panel of primary care providers that makes evidence-based recommendations on preventive services delivered in a primary care setting. There are two steps to this evidence-based decision-making process. The first step is collecting the relevant information and evaluating it through a systematic evidence review process, a step which is performed by the Evidence-based Practice Center (EPC) program. The second step is to make recommendations based on the strength of that evidence, and that is done by the USPSTF. Information about the USPSTF is available at our website, www.preventiveservices.ahrq.gov, which contains all of the evidence reports and all of the recommendations made by the USPSTF.

There is a generic analytic framework that is used by the USPSTF in evaluating any given screening topic. The USPSTF first specifies persons at risk and the tests to be used for screening. The EPC collects the evidence on the accuracy of the tests in detecting the target condition along with the benefits and harms of the tests. The next steps are to determine the treatment options for a person with that target condition and the benefits and adverse effects of the treatments. The primary interest is to clarify effect of treatments on health outcomes. When that information is not available, the effect on biomarkers (surrogate markers or intermediate outcomes) and the association of biomarkers in improving health outcomes is clarified. Therefore the analytic framework is a sequential chain of evidence that links a person at one end to improved outcomes at the other end. Sometimes there are trials that directly evaluate screening tests and morbidity/mortality outcomes, such as the trials for screening for breast cancer or for abdominal aortic aneurysm, eliminating the need to evaluate the evidence for intermediate steps. In the cardiovascular field, trials of lipid-lowering drugs have lead to

reduced morbidity and mortality, so there is no need to establish an association between that biomarker and a health outcome.

The USPSTF gives its recommendations in the form of letter grades, A through D, whenever the USPSTF finds at least fair quality evidence. Letter grade “A” is assigned when the USPSTF finds good evidence and that benefits substantially outweigh the harms. Letter grade “B” is assigned when the benefits do outweigh the harms, but not to the same degree as in letter grade “A”. The letter grade “C” is given when the benefits marginally outweigh harms. The letter grade “D” is given when there is at least fair evidence that either the service, whether it is the test or the treatment, is ineffective or that the harms outweigh the benefits. The letter grade “I” indicates that there is not enough evidence to make a recommendation, due to the lack of data, poor quality data or conflicting data.

As an example, the USPSTF set out to evaluate evidence for screening asymptomatic adults for early detection of abnormal lipids. Specifically, it wanted to better evaluate accuracy and potential harms of screening tests, the relevant interventions (diet therapy, exercise or pharmaceutical agents) and the evidence of improved health outcomes (reduced morbidity or mortality) and harms of interventions. As a result of evaluating the evidence, the USPSTF issued recommendations in 2001 supporting screening for lipid abnormalities and are planning to release an update.

Another example is screening for Hepatitis C virus (HCV) infection. Although it is not a cardiovascular disease example, it is a good example of inadequate evidence for a USPSTF recommendation. The first step is to identify risk factors for HCV infection. After the initial stratification of patients based on risk factors, the next step is to decide whether screening for HCV infection is warranted. Among the patients who are diagnosed with HCV infection and are eligible for antiviral therapy, the question is “what is the evidence that antiviral therapy leads to improvement in biomarkers (or surrogate outcomes)?” In this case, the biomarkers included viral titers and titers of liver enzymes. The next question is “what is the association of these biomarkers with decreased morbidity or mortality?” The USPSTF review determined that HCV screening tests are accurate and that there is a high prevalence of HCV infection in high-risk persons, maybe as high as 90% in intravenous drug users, but even as high as 50% in other high-risk groups. However, the natural history of HCV infection is not that clear. Only 10-20% of patients with HCV infection progress to cirrhosis of the liver over a fairly long period of time, 20-30 years. Only 30-40% of the infected patients are eligible for antiviral therapy. Therapy results in sustained virologic response in about 50% of the patients. What is not known is whether this improvement is associated with improved health outcomes: decreased morbidity or mortality from cirrhosis or hepatocellular carcinoma. Among the harms of therapy are high rates of adverse events. Nearly 60% of patients have adverse events and there is a fairly high rate of therapy discontinuation. Additionally, the therapy lasts for several months and is expensive. The USPSTF came to the conclusion that there is insufficient evidence to screen high-risk patients for HCV infection and recommended against screening for HCV infection in average-risk (or general) population.

There are some common questions of interest to all decision-makers:

- What are the relevant outcomes?
- Does the diagnostic test and subsequent therapy lead to improved patient outcomes?
- Do the benefits outweigh the harms (i.e. is there a net benefit)?
- What is the incremental benefit?
- What are the costs of the tests and therapy?
- Given the net health benefit and given the costs, is it worth it?

It is rare to find a condition where there are no available tests or therapies, so determining the added value (or incremental benefit) of a new diagnostic test or a new therapy is important. The next question is whether the benefits we see in clinical trials actually translate into real world experience. Economic aspects are also important. It is important to know the costs and cost-effectiveness of tests and treatments.

Moving to the topic of genomics tests, most of the gene based tests currently available have either no or poor quality health outcomes data. Second, the lab-developed tests are not regulated by the FDA. Two examples for breast cancer are the BRCA1 and BRCA2 tests, which apart from not being FDA regulated, are also patented. There is only one lab in the world that can perform these tests, unless the test is licensed to other labs. The Oncotype DX® is an example of a genomic test where several different genes are evaluated and used to predict risk for breast cancer recurrence to tailor therapy accordingly. That at least is the stated benefit of this test. There is no evidence yet in a prospective study that this actually works in practice. Another point about gene-based tests is that these are expensive. The final point is that the pathway of integrating genomic applications into clinical practice is very complex. There are different steps necessary to move from basic biomedical discovery to gene-based application, and then to routine clinical use. To begin with, a case is made that a given genetic variation is linked with a given disease or in some cases, it causes a disease. Then usually it is the role of pharmaceutical or biotechnology companies of using this knowledge to develop new diagnostic or therapeutic applications. The pathway for the drugs to clinical availability is better defined with Phase I, II and III trials and FDA approval. The pathway for diagnostics is not so well-defined, especially for gene-based tests. In order to transition from clinical availability to routine clinical use the first step is outcomes research. Outcomes research really covers a lot of territory since it could take the form of randomized controlled trials, observational studies, or case control studies. The next steps include the synthesis of evidence and the involvement of decision makers. One example of the synthesis of evidence is the process of systematic evidence reviews performed by the Evidence-Based Practice Center program. The Cochrane Collaboration is another example. These evidence reports are then used by decision makers including clinical guideline developers as well as insurers, and organizations such as CMS, that make coverage and reimbursement decisions.

A big challenge for AHRQ is when there is enough evidence that a new test or a new drug will work and improve health outcomes and we need to implement this new

information into practice. Implementation involves behavioral change, both at the provider level and at the patient level, and there are many different programs that are ongoing to support this effort including health information technology, computer decision aids, dissemination of information, financial incentives, etc. There are a large number of entities involved in this step before knowledge is translated into routine clinical use.

Finally, there are some other issues that deserve consideration. One of the issues is finding ways to enable the test developers to better assess the accuracy of the diagnostics. It would be desirable to have national repositories of samples, using standardized sample collection and reporting processes for use by researchers. The second issue for consideration, especially for genomics, is the use of standardized molecular data analysis methods. It is challenging to have different researchers using different techniques and coming up with results claiming to be the same. Additionally, we need to have clinical databases that link the diagnostic information with the drug information, to know whether or not the physicians decided to change therapy based on that information. We need to better understand what happened to patients, whether the patient was discharged early, whether they improved, deteriorated or the changed treatment had no effect. We also need to analyze other hospital data related to clinical process and workflow. We also need to improve the way we collect and analyze the data since most of these studies will be observational studies, which will have more risk of an error in the conclusions than randomized control trials. We need to find ways to improve our methodologies and reduce bias in observational studies. Finally, we need to consider different mechanisms and collaborations, public-private or private-private, to strengthen and improve our data collection and analysis.