

## Chapter 3

# Choosing the Important Outcomes for a Systematic Review of a Medical Test

Jodi B. Segal, M.D., M.P.H., Johns Hopkins University School of Medicine

### Abstract

In this chapter of the Evidence-based Practice Centers *Methods Guide for Medical Test Reviews*, we describe how the decision to use a medical test generates a broad range of outcomes, and suggest how each of these outcomes might be considered for inclusion in a systematic review. Awareness of these varied outcomes affects how a decisionmaker balances the benefits and risks of the test; therefore, a systematic review should present the evidence on their diversity. The key outcome categories include clinical management outcomes; direct health effects; emotional, social, cognitive, and behavioral responses to testing; legal and ethical outcomes; and costs. We describe the challenges of incorporating these outcomes in a systematic review, suggest a framework for generating potential outcomes for inclusion, and describe the role of stakeholders in choosing the outcomes for study. Finally, we give examples of systematic reviews that either included a range of outcomes or that might have done so. This chapter puts forward a set of key messages for systematic reviewers:

- Consider both the outcomes that are relevant to the process of testing and those relevant to the results of the test.
- Consider inclusion of outcomes in all five domains: clinical management effects; direct test effects; emotional, social, cognitive, and behavioral effects; legal and ethical effects; and costs.
- Consider to which group the outcomes of testing are most relevant.
- Given resource limitations, prioritize which outcomes to include. This decision depends on the needs of the stakeholder(s), who should be assisted in prioritizing the outcomes for inclusion.

### Introduction

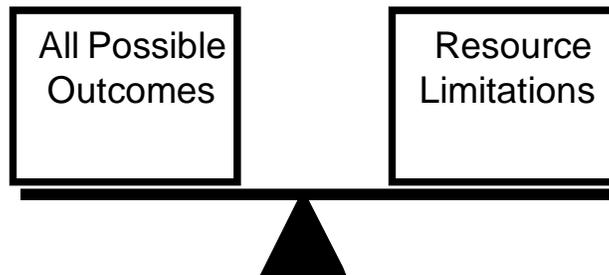
The Agency for Healthcare Research and Quality (AHRQ) requested production of a methods guide for comparative effectiveness reviews that specifically addresses the unique challenges of preparing a systematic review about the use of a medical test. This chapter describes the considerations to be taken into account when selecting the outcomes that will be included in a systematic review of a medical test. We describe the range of effects that medical tests have, and suggest how these outcomes from testing should be incorporated into a systematic review to make it maximally useful to those using the review.

We define “decision-relevant” outcomes as the outcomes that result from a testing encounter that may affect the decision to use the test. We consider a broad range of outcomes to illustrate how these may affect the balance of the benefits and risks of the test. The outcomes to be discussed in this chapter are those that are relevant to screening tests, diagnostic tests, and prognostic tests, although prognostic tests are also discussed in Chapter 11. We also address unique issues that might arise if the test in question is a genetic test, although genetic tests are explored in more detail in Chapter 10. We include a framework for generating potential outcomes for inclusion, and discuss the role of stakeholders in choosing the outcomes for study. Finally, we give examples of systematic reviews that either included a range of outcomes in the review or might have done so.

## Common Challenges

Investigators are tasked with choosing the outcomes to consider in a systematic review about a medical test. Resource limitations require judicious selection from among all possible outcomes, which necessitates setting priorities for the outcomes to include. If reviewers do not *explore* the full range of outcomes at the outset of the project, the likelihood of excluding important outcomes is high; the systematic review may miss outcomes relevant to the stakeholder(s). The balance of the benefits and harms from testing will be skewed by the absence of information about key outcomes. The consequence may be that recommendations based on the systematic review are inapt when the test is used in practice. Additionally, for tests that offer modest clinical gains over another test, information on additional outcomes, for example, costs or convenience, may be essential for making decisions based on the test results. However, we caution that if the initially broad range of outcomes is not carefully condensed, the quality of the review will be threatened by resource limitations (Figure 3–1).

Figure 3–1. Balance of outcomes against resources



Either misstep can result in a suboptimal review—the narrow review may be incomplete, and the broad review may be too superficial to provide meaningful insights.

## Principles for Addressing the Challenges and Recommended Approaches for Incorporating All Decision-Relevant Outcomes

We recommend a two-step approach for choosing the outcomes for inclusion in a review about a medical test. The first step is to catalog outcomes methodically, and the second is to solicit input from the stakeholder(s). Below is a description of a conceptual approach to identifying outcomes to ensure that relevant outcomes are not overlooked.

## Principle 1: Catalog outcomes methodically.

### Conceptual Approach To Identifying Outcomes

The preceding chapter described frameworks for designing systematic reviews about medical tests, including the PICOTs typology (i.e., population, intervention, comparisons, outcomes, timing, and setting). Here we present another framework specifically for thinking about the outcomes from using a test in a clinical setting. In this framework, outcomes are separated into those attributable to the testing process and those attributable to knowledge of the test results. In general, outcomes attributable to the testing process are direct effects of the test; outcomes attributable to the test results are more numerous and include the patient's response to the test results and how the patient and clinician act upon the results.

Bossuyt and McCaffery described a useful framework for thinking about patient outcomes attributable to medical testing.<sup>1</sup> They classified outcomes into three groups: (1) outcomes that result from clinical management based on the test results; (2) the direct health effects of testing; and (3) the patients' emotional, social, cognitive, and behavioral responses to testing. We extend this model by including two additional elements to arrive at five types of outcomes: (4) the legal and ethical effects of testing, which may or may not be a consideration depending on the test under consideration; and (5) the costs of testing. These five categories of outcomes can be associated with the testing process, or with the test result, or with both.

We suggest that the relative importance of these outcomes may differ substantially depending on whether the intention of the test is screening, diagnosis, or prognosis. (Table 3–1) To illustrate, the adverse emotional effects, and the legal and ethical outcomes of testing, might be more significant for medical tests used for screening than tests used for diagnosis, due to the high prevalence of false positive test results associated with many tests used for screening purposes. Additionally, screening tests are conducted in individuals who are without symptoms of the disease of interest, so any adverse or disruptive consequences of testing may be more pronounced. Mammography is a useful example, since the emotional reaction to a false positive test may be substantial. Correspondingly, the potential legal consequences of a false negative test are substantial, as a false negative test may lead to the filing of a malpractice suit. Missed diagnoses, in particular breast cancer diagnoses, are a large category of radiology-related malpractice suits.<sup>2</sup>

**Table 3–1. Outcomes that might be particularly consequential depending on type of medical test**

Outcomes	Screening Test	Diagnostic Test	Prognostic Test
Clinical management	+	+++	+++
Direct health effects	+	++	++
Emotional, social, cognitive, behavioral responses	+++	++	++
Legal and ethical	+++	++	++
Costs	++	++	++

Systematic reviewers should remember as well that a normal test result, that is a test that has correctly excluded the presence of disease, may be as affecting as a test that has made a diagnosis, and inclusion of outcomes resulting from a negative test may be important in the review. The primary studies of the medical test may have assessed behaviors and consequences after a normal test result, which may include additional testing when a diagnosis is sought or a change in behavior in response to a normal test result (e.g., less attention to healthy lifestyle or

possibly redoubled efforts at maintaining good health). These are all appropriate outcomes for consideration for inclusion in a systematic review.

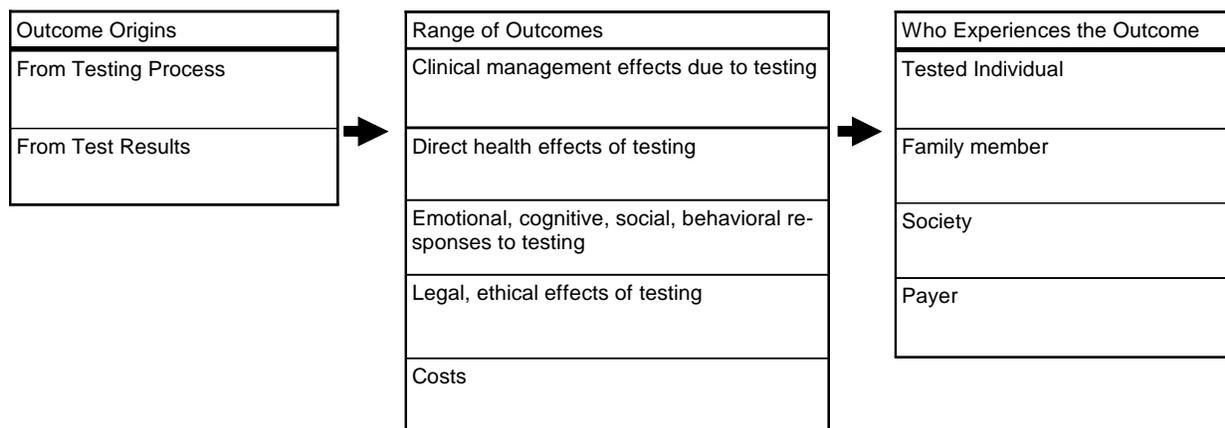
The impact of testing on clinical management is a more important consideration for reviewing diagnostic testing and less important for screening tests, where the clinical management may be quite removed from the screening step. A useful example of diagnostic testing is the use of computed tomography (CT) for detection of pulmonary embolism: a positive test will result in many months of anticoagulation therapy, an important clinical management consequence for the patient. Therefore, systematic reviews will ideally include primary literature that tests the clinical consequences resulting from the use of CT in this setting (rather than just the sensitivity and specificity and predictive values of the test). It is likely that the direct health effects of screening tests are less than in tests used for diagnosis and prognosis: screening tests are generally designed to be less invasive than tests used to make diagnoses in individuals suspected of having disease. An example is PAP testing for cervical cancer screening: there should be no direct health effects of this process.

The range of downstream activities that result from a test are also appropriate to consider for inclusion as outcomes. These may be particularly prominent in imaging tests where there is a high likelihood of identifying unexpected findings that necessitate further evaluation (e.g., unexpected adrenal masses seen during abdominal imaging), or in imaging tests that identify unexpected findings that worry the patient (e.g., degenerative spine changes seen on chest imaging). In selecting outcomes in these situations, one might consider the emotional and cognitive outcomes of unexpected findings, or the monetary costs of the downstream evaluation of incidentally identified abnormalities.

Additional cost outcomes might be considered if appropriate to the systematic review. In addition to the direct costs of the test, one might consider the downstream costs triggered by the results of the testing, which may include confirmatory testing following a positive result, treatment costs resulting from detecting disease, and costs for treatment of adverse effects of the testing (including direct harms of the test and downstream harms resulting from additional testing or treatment, or from evaluation of incidental findings.) Other costs to consider might be the costs to society from diversion of funds to testing and away from other services. As an example, one might include, in a systematic review of universal newborn screening, the impact of diverting funding away from other childhood programs such as vaccination.

In addition to consideration of the consequences of testing, we suggest that reviewers also consider an additional axis; namely, who experiences the outcome. The individual being tested is not the only one who can experience outcomes from the testing process. Outcomes may be experienced by family members, particularly in the case of testing an index person for heritable conditions. Outcomes may be experienced by the population *away* from which resources are diverted by a screening activity, e.g., widespread newborn screening which diverts resources away from population-based smoking cessation activities. Society as a whole may experience some outcomes, as when a test of an individual leads to a public health intervention, e.g. prophylactic antibiotics or quarantine after exposure to an infectious individual, or diversion of resources in order to pay for testing of other individuals. Payers are affected if they need to pay for a treatment of a newly diagnosed condition (Figure 3–2).

**Figure 3–2. Mapping outcomes to the testing process and to the test results**



In summary, a wide range of outcomes could be included in a systematic review of a test. We encourage investigators doing systematic literature reviews to think through this range of outcomes, considering the testing process, the test results, the range of associated outcomes, and the parties that may experience the outcomes. As we discuss below, these considerations may differ depending on the type of test under consideration, and will differ importantly by the specific test and the question being addressed by the systematic review.

## **Principle 2: Solicit input from stakeholders.**

### **Stakeholders’ Role in Defining the Outcomes and Guidance From the Reviewers**

Because the range of outcomes that a reviewer might include is broad, expecting such reviews to include “all possible outcomes” is unrealistic. The *AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews* (also referred to as the *General Methods Guide*) recommends that stakeholders be involved at several steps in the systematic review process.<sup>3</sup> We describe additional considerations regarding the role of stakeholders in reviews of medical tests, as their inputs are particularly relevant to the choice of outcomes for inclusion.

Little to no empiric evidence exists regarding what outcomes are most essential for inclusion in a systematic review. If the systematic reviewers knew that some outcomes are universally valued by users of reports, these would be routinely included. It is likely, however, that the choice of outcomes depends largely on the needs of stakeholders and how they intend to use the review. Clinicians and patients are frequently the primary users of the results of systematic reviews and therefore are important stakeholders in the outcomes selection process. An understanding of what evidence the patient or clinician needs in order to make a decision about use of the test is vital in selecting outcomes for inclusion in a review. Certainly the health effects of testing and the emotional or behavioral, social, or cognitive outcomes are directly relevant to the patient; a comprehensive review must include outcomes that are important to patients and would influence their use of a test.

To give an example of another type of stakeholder, the *Evaluation of Genomic Applications in Practice and Prevention* (EGAPP) group of the Centers for Disease Control and Prevention (CDC) has sponsored several EPC reports.<sup>4–6</sup> EGAPP uses these reports to generate guidelines that the CDC issues about genetic testing. EGAPP’s interests are broad; it aims to maximize the effectiveness of genetic testing at a societal level. Understandably, the outcomes that it considers

relevant are also broad, and range from the analytic validity of the test to the impact of the testing process on family members. When the possible outcomes for inclusion are many, the investigators have a responsibility to work with the stakeholder to refine the questions carefully so that the task can be accomplished.

Other stakeholders, like professional societies such as the American College of Physicians, may be most interested in evidence reports that can be used to generate recommendations or guidelines for practicing clinicians. Therefore, as stakeholders, they may be more focused on how clinical outcomes vary as a result of medical testing, and perhaps less interested in outcomes that may be more relevant to payers, such as cost-shifting to accommodate costs of testing and downstream costs.

Not infrequently, the primary users of systematic reviews are Federal agencies such as the Centers for Medicare & Medicaid Services (CMS). This agency is responsible for decisions regarding coverage of their beneficiaries' medical care, including medical tests. Therefore, CMS may specify that the outcome most relevant to its coverage decision is the analytic validity of the test, as it would not want to cover a test that inadequately identifies the condition of interest.

The researchers doing comprehensive systematic reviews have a role in helping stakeholders understand the breadth of outcomes. The researchers might assist stakeholders with mapping the range of outcomes depicted in Figure 3–2. This will allow the stakeholders to review the breadth of outcomes and characterize the outcomes as being more or less vital depending on the intended use of the review.

## **Illustrations of the Principles**

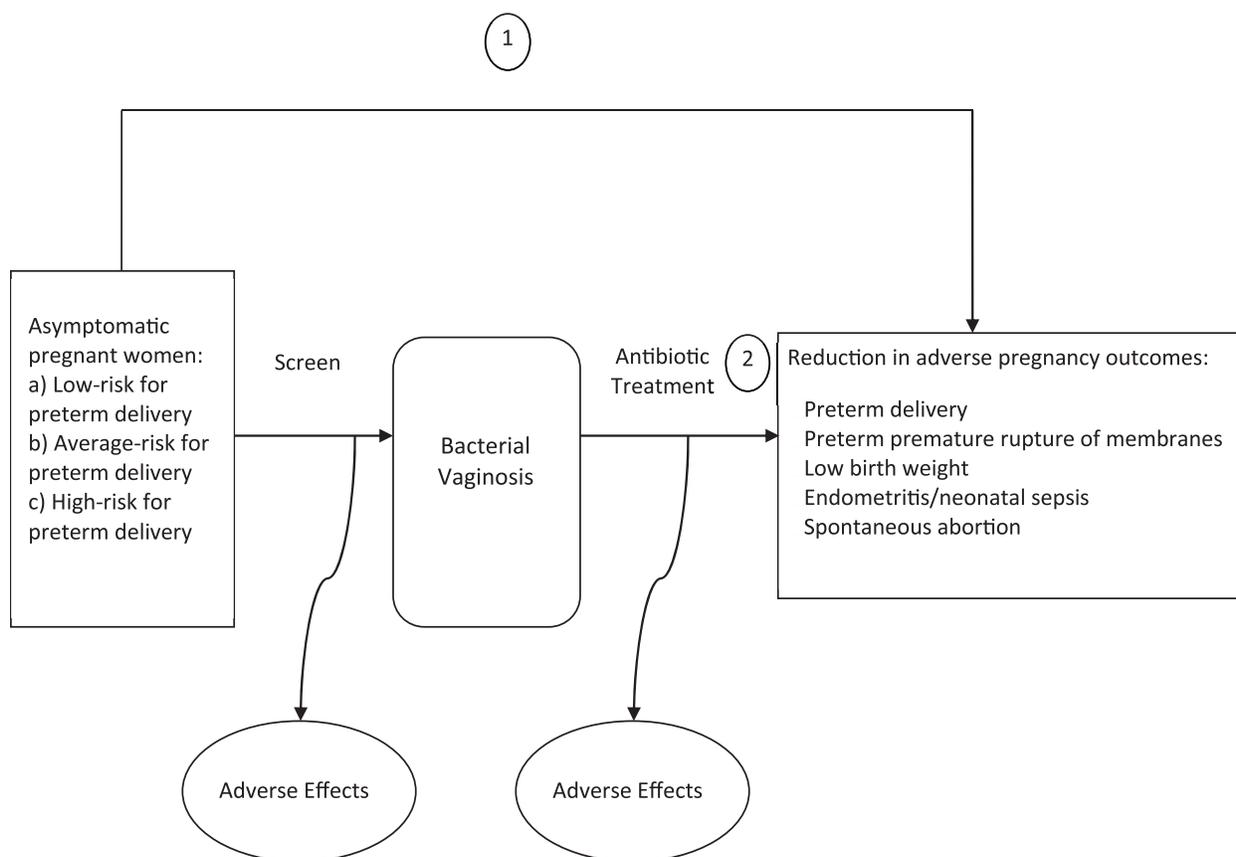
To explain these points in more detail, we use three examples: one each of a screening test, a diagnostic test, and a prognostic test. In discussing these examples, we consider both outcomes that result from the process of testing or are associated with the results of testing, and outcomes that affect the tested individual and others. We conclude with a discussion of additional considerations when the test is a genetic test.

### **Example of a Screening Test**

Screening tests are used to detect disease in asymptomatic individuals or individuals with unrecognized symptoms.<sup>7</sup> Screening tests should be able to separate individuals with the disease of interest from those without, and should be employed when there is a treatment available and where early treatment improves outcomes. The U.S. Preventive Services Task Force (USPSTF) develops recommendations for use of clinical preventive services in the United States. An EPC is sometimes tasked with preparing the supporting review of the evidence.<sup>8,9</sup> Other stakeholders have interest in screening tests as well, including professional organizations involved in guideline preparation for their practitioners; cases in point are recommendations made by the American College of Obstetrics and Gynecology regarding cervical cancer screening<sup>10</sup> and the American Cancer Society's recommendations for early cancer detection.<sup>11</sup>

To illustrate outcomes in a systematic review of a screening test, we present the example of a systematic review about screening for bacterial vaginosis in pregnant women.<sup>12</sup> This systematic review was first done for the USPSTF in 2001 and was later updated. Figure 3–3 depicts the analytic framework developed by the authors.

**Figure 3–3. Screening example: bacterial vaginosis**



*Clinical management effects.* The authors addressed whether screening for bacterial vaginosis during pregnancy in asymptomatic women reduces adverse pregnancy outcomes. They included a review of the clinical management effects that would result from antibiotic treatment based on screening results. These included adverse effects of therapies and beneficial effects—reduction in adverse pregnancy outcomes such as preterm delivery. The authors might also have explicitly included an outcome that examines whether the screening leads to receipt of antibiotic treatment—that is, whether screening leads to a change in clinical management. This would be a relevant intermediate outcome on the path between screening and the outcomes attributable to therapy.

*Direct test effects.* Appropriately, the authors of this review did not include outcomes that are a direct result of the testing process because direct test effects are unlikely in this example: a vaginal swab will not cause any injury. Similarly, the test does not confer any direct benefit either, except perhaps contact with clinicians.

*Emotional, social, cognitive, or behavioral effects.* The authors might have also looked at the emotional, social, cognitive, or behavioral effects of the screening process or the screening test results. It might have been appropriate to consider outcomes that are associated with screening but are not the result of antibiotic therapy. Consideration might have been given to the effects of testing positive for bacterial vaginosis, such as emotional responses to a diagnosis of infection, leading to either healthier or riskier prenatal activities, or maternal worry as an outcome.

As with any measure, the systematic review team might require that the instrument used to measure emotional response be a validated and appropriate instrument.

*Legal and ethical effect of testing.* Although specifying ethical issues in screening for bacterial vaginosis (which is not a sexually transmitted infection) may seem unnecessary, bacterial vaginosis testing may be done as part of an infectious disease screening for reportable diseases such as syphilis or HIV. Therefore, a review of the effects of testing should consider whether the test being reviewed might be administered with concurrent screening tests that could themselves raise ethical issues.

*Costs of the test.* The authors of this review did not consider the costs of the test to the patient as an outcome. Widespread initiation of screening programs, such as on a population level, may have profound cost implications.

The authors of this review considered the effects of screening on the mother and on the fetus or infant. However, they might have also considered other relevant parties; these might include the mother's partner and society, as antibiotic resistance is a conceivable outcome from widespread testing and treatment of bacterial vaginosis.

### **Example of a Diagnostic Test**

We differentiate diagnostic tests from screening tests largely by the population being tested. Whereas a diagnostic test is applied to confirm or refute disease in a symptomatic person, a screening test is used in an asymptomatic or pre-symptomatic person. The USPSTF mostly makes recommendations about screening tests that may be used in the general population; other organizations are more concerned with ensuring safe use of diagnostic tests in patient populations. Payers are also interested in optimizing use of diagnostic tests, as many are costly.

We discuss a review that addressed the diagnostic value of 64-slice computed tomography (CT) in comparison to conventional coronary angiography.<sup>13</sup> Stating that their review concerned the "accuracy" of CT, the authors aimed to assess whether 64-slice CT angiography might replace some coronary angiography for diagnosis and assessment of coronary artery disease. A broader review may consider the effectiveness of CT angiography, and the investigators would consider the full range of outcomes as below.

*Clinical management effects.* Numerous clinical management effects might follow testing for coronary artery disease with CT. The authors of the review focused exclusively on detection of occluded coronary arteries and not on any downstream outcomes from their identification. Individuals diagnosed with coronary artery disease are subjected to many clinical management changes; these include medications, recommendations for interventions such as angioplasty or bypass surgery, and recommendations for lifestyle changes; each of these changes has associated benefits and harms. All of these may be appropriate outcomes to include in evaluating a diagnostic test. If one test under consideration identifies more coronary artery disease than another, this difference will be reflected in clinical management and the consequences of the management choices.

Other conceivable clinical management effects relate to the impact of testing on other health maintenance activities. For example, a patient might defer other necessary testing (e.g., bone densitometry) to proceed with the CT. We would expect, however, that this would also be the case in the comparison arm. Family members may be affected as well by testing; for instance,

they may be called upon to assist the diagnosed patient with future appointments, which may necessitate time away from work and cause emotional stress.

*Direct test effects.* The test under consideration is a radiographic test. It confers no direct benefit itself (unlike the comparison procedure in which an intervention can be performed at the time of conventional diagnostic angiography). The testing process poses potential harms, including allergic reaction to the intravenous contrast material, renal failure from the contrast material, and radiation exposure. These are all outcomes that could be considered for inclusion. In this example, the comparison test carries comparable or greater risks.

*Emotional, social, cognitive, or behavioral effects.* The testing process itself is unlikely to have significant emotional consequences, as it is not an invasive test and is generally comfortable for the tested individual. The results of testing could indeed have emotional or behavioral consequences. An individual diagnosed with coronary disease might alter his or her lifestyle to reduce disease progression. On the other hand, an individual might become depressed by the results and engage in less self-care or riskier behavior. These behavioral effects are likely to affect the family members as well. However, in this example the emotional or behavioral effects are expected to be similar for both CT and conventional angiography and therefore may not be relevant for this particular review. In contrast, they would be relevant outcomes if CT angiography were being compared with no testing.

*Legal and ethical effects of testing.* Testing could have legal consequences if the tested individual is in a profession that requires disclosure of health threats for the safety of the public; this might arise if, e.g., the tested person is an airline pilot. However again, this outcome is not expected to differ between CT and conventional angiography.

*Costs of the test.* The relative costs of the two tests to the insurer and the patient, and the costs of diverting equipment away from other uses, could also be of interest to some stakeholders.

### **Outcomes Unique to Prognostic Tests**

A prognostic test is a test used in individuals with known disease to predict outcomes. The procedure itself may be identical to a procedure that is used as a screening test or a medical test, but the results are applied with a different purpose. Given these differences, additional considerations for outcomes should be included in reviews. For example, consider the use of spirometry for predicting prognosis in individuals with chronic obstructive pulmonary disease (COPD). The test is commonly used to make the diagnosis of COPD and to monitor response to treatment, but the question has been raised as to whether it might also predict survival. In 2005, the Minnesota EPC did a systematic review of this topic on behalf of the American Thoracic Society, American College of Physicians, American Academy of Family Physicians, and American Academy of Pediatrics.<sup>14</sup> The discussion below focuses on one of their key questions: whether prediction of prognosis with spirometry, with or without clinical indicators, is more accurate than prediction based on clinical indicators alone. The sponsoring organizations were interested in predicting patients' survival free of premature death and disability.

*Clinical management effects.* The results from prognostic testing will have effects on clinical management. Although the prognoses for some diseases are minimally modifiable with current treatments, most prognostic information can be used to alter the course of treatment. In this

example, spirometry may suggest a high likelihood of progressing to respiratory failure and prompt interventions to avert this (e.g., pulmonary rehabilitation efforts, changes in medication, avoidance of some exposures). Conversely, the prognostic information may be used to make decisions regarding other interventions. If the likelihood of dying from respiratory failure is high, patients and their physicians may choose to refrain from colonoscopy and other screening procedures from which the patient is unlikely to benefit. Similarly, treatments of other conditions may be of less interest if life expectancy is short.

*Direct test effects.* Spirometry has few direct test effects, although patients can have adverse reactions to testing, particularly if challenged with methacholine as part of the test. In general, it is unlikely that tests used for prognosis are more or less likely to have direct test effects than tests used for other purposes.

*Emotional, social, cognitive, or behavioral effects.* We doubt that many emotional or cognitive effects would arise in response to the testing process. Spirometry is a noninvasive test that most patients tolerate well. Emotional effects in response to the results of testing are possible; emotional effects could even be more pronounced for prognostic tests than for screening or medical tests if the test yields more specific information about mortality risk than is usual from a diagnostic test. There could be a range of effects on behavior including efforts to alter prognosis, like quitting smoking. Test results with prognostic information would be expected to affect family members as well.

*Legal and ethical effects of testing.* Results of tests that provide prognostic information could have legal outcomes, too, especially if the tested individual acts in ways that belie the information he has received (e.g., entering into a contract or relationship that he is unlikely to fulfill). In this present example, it is unlikely that the prognostic information from spirometry would actually raise these issues, but in other cases, such as a test that demonstrates widely metastatic cancer, this could be an issue. These legal and ethical effects of testing may reach beyond the tested individual and affect society if many individuals have substantial concealed information that influences their actions.

*Costs of the test.* The relative costs of the test to the insurer and the patient, relative to the costs of collecting information from a history and physical examination, may all be of interest to stakeholders.

### **If the Test Is a Genetic Test**

Chapter 10 of this guide describes in detail unique issues regarding evaluation of genetic tests. With respect to relevant outcomes, we note a few considerations here. Most prominent is the effect on family members. Genetic information about the tested individual has direct bearing on family members who share genes. This may affect emotional and behavioral outcomes, and ethical outcomes, if family members feel pressured to proceed with testing to provide better information for the rest of the family. A second issue is possible impact on health insurance eligibility. Recent legislation in the United States prohibits the use of genetic test results to exclude an individual from health insurance coverage, making this less a relevant outcome than in the past. This policy varies worldwide, however, and may be a relevant consideration in some countries.

## Summary

In specifying and setting priorities for outcomes to address in their systematic reviews, investigators should:

- Consider both outcomes that are relevant to the process of testing and those that are relevant to the results of the test.
- Consider inclusion of outcomes in all five domains: clinical management effects, direct test effects; emotional, social, cognitive and behavioral effects; legal and ethical effects; and costs.
- Consider to which group the outcomes of testing are most relevant.
- Given resource limitations, prioritize which outcomes to include. This decision depends on the needs of the stakeholder(s), who should be assisted in prioritizing the outcomes for inclusion.

## References

1. Bossuyt PM, McCaffery K. Additional patient outcomes and pathways in evaluations of testing. *Med Decis Making* 2009;29:E30-E38.
2. Berlin L, Berlin JW. Malpractice and radiologists in Cook County, IL: trends in 20 years of litigation. *AJR Am J Roentgenol* 1995;165:781-788.
3. Agency for Healthcare Research and Quality. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville, MD: Agency for Healthcare Research and Quality; 2008–. <http://www.ncbi.nlm.nih.gov/books/NBK47095>. Accessed April 25, 2012.
4. Matchar DB, Thakur ME, Grossman I, et al. Testing for Cytochrome P450 Polymorphisms in Adults with Non-Psychotic Depression Treated With Selective Serotonin Reuptake Inhibitors (SSRIs). Evidence Report/Technology Assessment No. 146. AHRQ Publication No. 07-E002. Rockville, MD: Agency for Healthcare Research and Quality; January 2007. <http://www.ahrq.gov/downloads/pub/evidence/pdf/cyp450/cyp450.pdf> Accessed April 25, 2012.
5. Bonis PA, Trikalinos TA, Chung M, et al. Hereditary Nonpolyposis Colorectal Cancer: Diagnostic Strategies and Their Implications. Evidence Report/Technology Assessment No. 150. AHRQ Publication No. 07-E008. Rockville, MD: Agency for Healthcare Research and Quality; May 2001. <http://www.ncbi.nlm.nih.gov/books/NBK38285>. Accessed April 25, 2012.
6. Segal JB, Brotman DJ, Emadi A, et al. Outcomes of Genetic Testing in Adults with a History of Venous Thromboembolism. Evidence Report/Technology Assessment No. 180. AHRQ Publication No. 09-E011. Rockville, MD: Agency for Healthcare Research and Quality; June 2009. <http://www.ncbi.nlm.nih.gov/books/NBK44591> Accessed April 25, 2012
7. Wilson JM, Jungner YG. Principles and practice of mass screening for disease. *WHO Chronicle* 1968;22:473.
8. Hillier TA, Vesco KK, Pedula KL, Beil TL, Whitlock EP, Pettitt DJ. Screening for gestational diabetes mellitus: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2008;148:766-775.
9. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2008;149:638-658.
10. Waxman AG. Guidelines for cervical cancer screening: history and scientific rationale. *Clin Obstet Gynecol* 2005;48:77-97.
11. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2008: a review of current American Cancer Society guidelines and cancer screening issues. *CA Cancer J Clin* 2008;58:161-179.

12. Nygren P, Fu R, Freeman M, Bougatsos C, Guise JM. Screening and Treatment for Bacterial Vaginosis in Pregnancy: Systematic Review to Update the 2001 U.S. Preventive Services Task Force Recommendation. Evidence Synthesis No. 57. AHRQ Publication No. 08-05106-EF-1. Rockville, Rockville, MD: Agency for Healthcare Research and Quality; January 2008.
13. Mowatt G, Cook JA, Hillis GS et al. 64-Slice computed tomography angiography in the diagnosis and assessment of coronary artery disease: systematic review and meta-analysis. *Heart* 2008;94:1386-1393.
14. Wilt TJ, Niewoehner D, Kim C et al. Use of spirometry for case finding, diagnosis, and management of chronic obstructive pulmonary disease (COPD). *Evid Rep Technol Assess (Summ)* 2005;1-7.

**Funding:** Funded by the Agency for Health Care Research and Quality (AHRQ) under the Effective Health Care Program.

**Disclaimer:** The findings and conclusions expressed here are those of the authors and do not necessarily represent the views of AHRQ. Therefore, no statement should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

**Public domain notice:** This document is in the public domain and may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the specific permission of copyright holders.

**Accessibility:** Persons using assistive technology may not be able to fully access information in this report. For assistance contact [EffectiveHealthCare@ahrq.hhs.gov](mailto:EffectiveHealthCare@ahrq.hhs.gov).

**Conflict of interest:** The author has no affiliations or financial involvement that conflicts with the information presented in this chapter.

**Corresponding author:** Jodi B. Segal, M.D, M.P.H., 1830 E. Monument St., Room 8047, Baltimore, MD 21287. Phone (410) 955-9866; Fax (410) 502-6952; email [jsegal@jhmi.edu](mailto:jsegal@jhmi.edu)

**Suggested citation:** Segal JB. Choosing the important outcomes for a systematic review of a medical test. AHRQ Publication No. 12-EHC075-EF. Chapter 3 of Methods Guide for Medical Test Reviews (AHRQ Publication No. 12-EHC017). Rockville, MD: Agency for Healthcare Research and Quality; June 2012. [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm). Also published in a special supplement to the Journal of General Internal Medicine, July 2012.