HEALTH TECHNOLOGY ASSESSMENT EVIDENCE CRITERIA: WHAT TYPES OF EVIDENCE SHOULD BE PRESENTED FOR PRODUCTS USED TO SCREEN FOR DISEASE IN THE UNITED STATES?

Nancy Mari McGee, JD, MPH

A dissertation submitted to the faculty of the University of North Carolina Chapel Hill in partial fulfillment of the requirements of the degree of Doctor of Public Health in the Gillings School of Global Public Health

Chapel Hill
2011

Approved by:

John Paul, PhD – Chairman

Margaret Dardess, JD, PhD

Sheila K. Fifer, PhD

Norrie A. Thomas, PhD

Morris Weinberger, PhD

©2011
Nancy Mari McGee
ALL RIGHTS RESERVED

ABSTRACT

Health Technology Assessment Evidence Criteria:
What Types of Evidence Should be Presented for Products Used to Screen for Disease in the United States?
(Under the direction of John Paul, PhD)

New technologies can be tools of innovative change in healthcare. They can be associated with improved treatment options, quality of care for patients, and cost savings. Distinguishing valuable technologies from those that offer added costs with no improvements in outcomes is the art of technology assessment. A key function of that art involves selecting those patients, conditions, providers and settings in which a technology may offer improvements over current care. With recent scientific discoveries such as the mapping of the human genome, development of genetic marker tests, and the growing interest in stem cell technologies, innovation is far ahead of any type of assessment that is currently used to establish which technologies should be made accessible to patients.

Screening technologies are on the forefront of innovation and may have a dramatic impact on the care of patients in terms of identifying disease and appropriate treatment options at an early stage. As a result, screening technologies are of key interest to health technology assessment (HTA) agencies in the United States and abroad. Similarly, because screening technologies are developing quickly and are believed to have the potential to make a significant change in patient care, it is important to develop a robust level of HTA criteria to evaluate these new technologies and determine which

technologies should be integrated into the practice of medicine and made accessible to patients.

Findings from this study indicate that while technology assessment organizations do have standard sets of criteria to evaluate products that are therapeutic, the assessment and level of evidence used to evaluate screening technologies are less clear. The objective of this research study is to evaluate existing technology assessment standards for screening technologies in order to establish a best practice that may be implemented by US technology assessment organizations to broaden the criteria used in assessments for screening products. The results of this study indicate that the best practices should include criteria to: support screening reliability, sensitivity and specificity; evaluate data to identify appropriate patient populations; reference to the natural course of the disease; consider ethical implications; and the impact of cost.

To:
Mari McGee Truumaa and Dr. Aare Truumaa

ACKNOWLEDGEMENTS

A doctoral dissertation may appear to be a solitary work. However, to complete an endeavor of this magnitude requires a network of support, and I am indebted to many people. I am most especially grateful to Dr. Michael Pasyanos for his encouragement, thoughtful input, and unconditional love.

I am grateful to my colleagues. In particular, Dr. Lee Eisenberg, MPH, is a mentor and friend who encourages and inspires me. Dr. Sheila Fifer and Dr. Norrie Thomas both provide me with unwavering support and counsel. J.D. Kleinke helps me to think broadly, push to the edge, and challenge the norm.

For this study, Dr. Clifford Goodman provided personal guidance and recommendations. In addition, I would like to give my most sincere thanks to Susanne Fausing Machacz, MPH and Katie Drasser, MSPH. This study benefited from their passion for health policy and dedication to patient access issues.

Finally, I would like to acknowledge two important people who are no longer with us. Dr. Ronald Dei Cas, MPH served as a mentor and close friend for many years. Ron sent an enthusiastic letter of recommendation to the University of North Carolina to support my admission to the doctoral program and was excited that I was pursuing this study. My father, R. John McGee, was a pharmacist in San Diego for over 30 years. His dedication to the physicians and patients that he served in the community sparked a curiosity in me about healthcare that has shaped my career.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

New technologies can be tools of innovative change in healthcare. They can be associated with improved treatment options, improved quality of care for patients, and cost savings. Distinguishing valuable technologies from those that offer added costs with insufficient improvements in outcomes to offset added marginal costs is the art of technology assessment. A key function of that art also involves selecting those patients, conditions, providers and settings in which a technology may offer improvements over current care. With recent scientific discoveries such as the mapping of the human genome, development of genetic marker tests, and the growing interest in stem cell technologies, innovation is far ahead of any type of assessment that is currently used to establish which technologies should be made accessible to patients.

Screening technologies offer a useful case study in where and how technology assessment as currently practiced stumbles when applied to what might be called "non-traditional" innovations in medicine. Screening devices are on the forefront of innovation and may have a dramatic impact on the care of patients in terms of identifying disease and appropriate treatment options at an early stage. As a result, screening technologies are of key interest to health technology assessment (HTA) agencies in the United States and abroad. Similarly, because screening technologies are developing quickly and are believed to have the potential to make a significant change in patient care, it will be important to develop a robust level of HTA criteria to evaluate these new

technologies and determine which technologies should be integrated into the practice of medicine and made accessible to patients during their course of care.

Health Technology Assessment

Health technology assessment (HTA) is the applied arm of evidence-based medicine that is often used by payers and providers to determine patient access to new technology. Traditionally, technology assessments have been applied to new healthcare technologies that provide a therapeutic outcome in patients. Evaluating screening technologies is an area of HTA that is evolving. One complicating factor is that outcomes evidence is particularly difficult and expensive to obtain in evaluating interventions for screening tests and procedures. As this is a new area in healthcare and may provide a pathway to discovering disease and evaluating treatment options earlier in the clinical pathway, it is particularly important to understand the level of evidence required for screening products to successfully meet the standards of the HTA.

The purpose of HTA is to review new technologies, by examining existing evidence, in order to understand if the technology is effective as compared to existing methods. In addition, the purpose of HTA is to help healthcare stakeholders, such as payers, ensure that their choices about whether or not to cover a technology is consistent with their stated goals and obligations to their beneficiaries.

Technology assessment is a concept that first arose in the mid-1960s from a burgeoning appreciation of the critical role of technology in society. In addition to the positive aspects of new technology, there was also a growing realization for the potential of unintended, possibly harmful, consequences of adopting technology before

understanding whether or not it was safe.[1] In this sense, technology assessment "was conceived as a way to identify the desirable first-order, intended effects of technologies as well as the higher-order, unintended social, economic and environmental effects."[2]

There is no single standard for conducting HTA in the US. Currently, assessments are conducted by diverse groups of stakeholders that apply analytic frameworks that draw on a variety of methods for a variety of objectives. Examples of types of organizations that conduct and employ HTA include: regulatory agencies, government and private payers, managed care organizations, health professions organizations, group purchasing organizations, and investors.[3] Depending on the frameworks and methods that are employed by the reviewing agency, the findings of a HTA conducted by separate organizations may be different and conflicting for the same technology. This result may occur because stakeholder preferences drive the types and levels of evidence used in the technology assessments. For example, in Europe, HTA conducted by government payers employ cost as an evidence measure;[4] while in the US, cost has typically not been included in assessments conducted by public payers.

Screening Technologies

Screening technologies may be able to help identify whether or not a patient has a specific disease in order to inform treatment decisions before symptoms occur. In

---

[1] Goodman, C. HTA 101 Introduction to Health Technology Assessment. 2004. http://www.nlm.nih.gov/nichsr/hta101/ta10103.html#Heading4 Accessed January 25, 2009.

[2] *Ibid*.

[3] *Ibid*. http://www.nlm.nih.gov/nichsr/hta101/ta10104.html#Heading7 Accessed January 25, 2009.

[4] NICE: Our Guidance Sets the Standard for Good Healthcare; NHS National Institute for Health and Clinical Evidence. Page 9. 30 December 2008. http://www.nice.org.uk/aboutnice/?domedia=1&mid=EE5AA72F-19B9-E0B5-D4215C860E77FD2E Accessed February 22, 2009.

addition, screening might also identify risk factors beyond phenotype risks. For some diseases earlier diagnosis is associated with greater probability of cure or successful management – for others not. This varies depending upon the treatments that are currently available. For example, when treatments are available for a particular diagnosis a patient's life might be improved by reducing the burden of a treatment by finding the disease before it progresses. Similarly, when treatments are available and can be pursued, the patient's emotional burden of having an illness might be reduced. However, conversely, if a treatment does not exist, screening for disease and finding disease might increase the emotional burden on the patient.

There are confounders related to screening technologies. There is not consensus as to whether the discovery of a disease in an early phase may ultimately cost the insurer more or less to treat the patient. In addition, the ultimate need for screening might be called into question in cases in which defining when a risk factor indicates that treatment is needed versus identifying a pre-condition that may not progress into disease. Another controversy concerns the length of the observation period. Is a screening useful in defining a meaningful proportion of patients found to be at high risk for developing symptoms or negative health outcomes after 1 year, 10 or 20 or only after 30?

Regardless of the confounders, a significant amount of private sector innovation is occurring to develop new technologies that can identify a disease early. As such, it is unclear how far down the clinical pathway one must go in order to validate that a screening product actually helped the patient. Because these interventions typically do not cure the condition and may not be directly linked to higher rates of cures or improved management, their outcomes measurements address metrics such as earlier diagnosis and

identification symptoms in order to improve patient reported outcomes and/or delayed onset of disease.

Objective of Research Study

While technology assessment organizations do have standard sets of criteria to evaluate products that are therapeutic, the assessment and level of evidence used to evaluate screening technologies have not caught up with the development of innovative new products to screen for disease. Thus, the objective of this research study is to evaluate existing technology assessment standards for screening technologies in order to create a best practice to expand the types of criteria that might be evaluated during an HTA of a screening technology. These best practices will guide the process of evaluating new criteria and may be implemented by U.S. technology assessment organizations in the future.

Research Study Deliverables

In order to accomplish this objective, the following research steps will be conducted and deliverables created:

1. Survey identified selected HTA agencies to at least one publicly available assessments, and related materials, if available, for the same screening technology/product.

   Deliverable: Summary of searched HTA websites to identify relevant screening technology/product.

2. Create an evidence table that compares and displays the similarities and differences between the HTAs.

   Deliverable: Evidence table with written analysis.

3. Develop an interview guide based on the analysis.

   Deliverable: Draft interview guide for Dissertation Committee review, IRB clearance, and comment.

4. Schedule and conduct telephone interviews with 8-10 HTA decision-makers. Decision-makers will be identified from the research of the specific screening technology/product, from suggestions from Dissertation Committee members, and individuals identified during the research process.

   Deliverable: Summary and analysis of the interview results.

5. Synthesis of data and information from the evidence table and interviews.

   Deliverable: Suggested criteria for screening technologies and recommendation for implementation.

Plan for Change

The plan for change resulting from this research will be the creation of best practice that health technology assessment organizations may consider in assessing screening products in the future. The objective of the best practice will be to refine traditional HTA criteria to include different types of evidence that are more relevant to screening products. The ability to take non-traditional types of evidence into consideration during a screening product HTA will focus assessments as to whether or not a particular screening technology might help a specific patient population. As a result, more screening products that are appropriate for specific patient populations will be available.

Refined criteria for screening products will help organizations that use HTA better understand which technologies meet their stated organizational goals and

objectives so that they may make intentional choices about which screening products bring the most value to their patient populations. The best practices will help to direct better private sector investment in technologies that deliver improved health outcomes. Patients, manufacturers, and physicians will charge payers and HTA organizations to consider the refinement of criteria for screening products via this best practice so that specific patient populations are not uniformly denied access to screening products which may be highly effective.

Chapter 2: Literature Review

This section describes the methods and results of the literature review conducted for this study. As a preliminary step, definitions were researched to provide overall guidance to the topics to be researched. Three steps were then performed for the literature examined: (1) search; (2) selection – inclusion and exclusion criteria; and (3) review.

Definitions

Technology Assessment[5] – Technology assessment (TA) is a category of studies, intended to provide decision makers with information about the possible impacts and consequences of a new technology or a significant change in an old technology. TA is concerned with both direct and indirect or secondary consequences, both benefits and disadvantages, and with mapping the uncertainties involved in any government or private use or transfer of a technology. TA provides decision makers with an ordered set of analyzed policy options, and an understanding of their implications for the economy, the environment, and the social, political, and legal processes and institutions of society.[6]

Technology assessment is a form of policy research that examines short- and long-term social consequences (for example, societal, economic, ethical, and legal) of the

---

[5] Goodman, C. HTA 101 Introduction to Health Technology Assessment. 2004. "Some Definitions of TA and HTA". http://www.nlm.nih.gov/nichsr/hta101/ta10103.html#Heading4 Accessed August 2, 2009.

[6] Coates & Jarratt, Inc. *Course Workbook: Technology Assessment. Anticipating the Consequences of Technological Choices*. 1992. Washington, DC.

application of technology. The goal of technology assessment is to provide policy-makers with information on policy alternatives.[7]

Health Technology Assessment[8] (United States) – Health technology assessment is a structured analysis of a health technology, a set of related technologies, or a technology-related issue that is performed for the purpose of providing input to a policy decision.[9]

Health Technology Assessment[10] (United Kingdom) – Health technology assessment considers the effectiveness, appropriateness and cost of technologies. It does this by asking four fundamental questions: Does the technology work, for whom, at what cost, and how does it compare with alternatives?[11]

Screening – performance of a test in an asymptomatic population with the purpose of reducing morbidity and/or mortality from disease.[12]

Screening Test – A screening test is any testing procedure designed to separate people or objects according to a fixed characteristic or property, with the intention of detecting

---

[7] Banta HD, Luce BR. (1993). *Health Care Technology and Its Assessment: An International Perspective*. New York, NY: Oxford University Press.

[8] Goodman, C. HTA 101 Introduction to Health Technology Assessment. 2004. "Some Definitions of TA and HTA". http://www.nlm.nih.gov/nichsr/hta101/ta10103.html#Heading4 Accessed August 2, 2009.

[9] US Congress, Office of Technology Assessment. *Protecting Privacy in Computerized Medical Information*. Washington, DC: US Government Printing Office; 1994.

[10] Goodman, C. HTA 101 Introduction to Health Technology Assessment. 2004. "Some Definitions of TA and HTA". http://www.nlm.nih.gov/nichsr/hta101/ta10103.html#Heading4 Accessed August 2, 2009.

[11] UK National Health Service R&D Health Technology Assessment Programme, 2003. http://www.ncchta.org/about/whatishta.shtml. Accessed August 2, 2009.

[12] Sunshine, JH, et al. Technology assessment for radiologists. *Radiology*, 2004; 230:309-314.

early evidence of disease.[13] For purposes of this research study, a screening technology is a type of screening test or procedure used to detect early evidence of disease or to identify patient population segments who should receive diagnostic tests.

Diagnosis – "the process of determining by examination the nature and circumstances of a diseased condition."[14]

Diagnostic – is a tool, such as a test, used to determine a medical diagnosis.

QALY – A Quality Adjusted Life Year (QALY) is a calculation that provides an idea of how many extra months or years of life of a reasonable quality a person might gain as a result of treatment.[15] It is used to determine the monetary value of a particular medical treatment.

Comparative Effectiveness Research – Comparative effectiveness research (CER) "is the 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 healthcare at both the individual and population levels."[16]

---

[13] Stedman's Online Medical Dictionary. http://www.stedmans.com/section.cfm/45. Accessed: August 2, 2009.

[14] Definition "diagnosis". http://dictionary.reference.com/browse/diagnosis. Accessed July 1, 2011.

[15] Measuring Effectiveness and Cost Effectiveness: the QALY. National Institute for Health and Clinical Excellence. http://www.nice.org.uk/newsroom/features/measuringeffectivenessandcosteffectivenesstheqaly.jsp. Accessed July 1, 2011.

[16] Initial National Priorities for Comparative Effectiveness Research, Committee on Comparative Effectiveness Research, Prioritization Board on Health Care Services, page S-9. Institute of Medicine of the National Academies. http://www.nap.edu/catalog.12648.html. Accessed December 6, 2009.

Literature Search

Healthcare technology assessments intersect areas related to both healthcare delivery and policy. For purposes of this literature review, databases and websites were used to learn how researchers and policy makers assess technologies related to screening products.

PubMed was the database resource used to identify peer-reviewed literature for this review. Key word searches were utilized to narrow in on thematic issues around healthcare technology assessment, diagnostics, screening, and criteria. Diagnostics are included in the search criteria in order to capture references that might include both diagnostic and screening products. The initial key words that were used for the literature search are shown, below, in Table 1.

Table 1. Initial Key Words

Healthcare	AND	Technology A/E	AND	C/D/S	AND	Test
Healthcare		Technology Assessment		Criteria		
OR		OR		OR		
Healthcare		Technology Evaluation		Criteria		
OR		OR		OR		
Healthcare		Technology Assessment		Diagnostic		
OR		OR		OR		
Healthcare		Technology Assessment		Screening		
OR		OR		OR		
Healthcare		Technology Evaluation		Diagnostic		
OR		OR		OR		
Healthcare		Technology Evaluation		Screening		
OR		OR		OR		
Healthcare		Technology Assessment		Diagnostic		Test
OR		OR		OR		OR
Healthcare		Technology Assessment		Screening		Test
OR		OR		OR		OR
Healthcare		Technology Evaluation		Diagnostic		Test
OR		OR		OR		OR
Healthcare		Technology Evaluation		Screening		Test

An initial search on PubMed was conducted using the search terms yielded the preliminary search results set forth, below, in Table 2.

Table 2. PubMed Initial Key Word Search Results

Key Words	Search Results
Healthcare Technology Assessment Criteria	1,287
Healthcare Technology Evaluation Criteria	1,571
Healthcare Technology Assessment Diagnostic	1,437
Healthcare Technology Assessment Screening	776
Healthcare Technology Evaluation Diagnostic	1,789
Healthcare Technology Evaluation Screening	996
Healthcare Technology Assessment Diagnostic Test	47
Healthcare Technology Assessment Screening Test	217
Healthcare Technology Evaluation Diagnostic Test	61
Healthcare Technology Evaluation Screening Test	327

The websites researched include institutions and agencies that have provided guidance and influence regarding HTA and the development of HTA criteria and standards of evidence. These websites are set forth in Table 3.

Table 3. HTA Websites

Institution/Agency	Website
Centers for Medicare and Medicaid Services	http://www.cms.hhs.gov/default.asp?
Agency for Healthcare Research and Quality	http://www.ahrq.gov/
Blue Cross Blue Shield Association Technology Evaluation Center	http://www.bcbs.com/blueresources/tec/
U.S. Preventative Services Task Force	http://www.ahrq.gov/CLINIC/uspstfix.htm
National Library of Medicine – National Institutes of Health	http://www.nlm.nih.gov/
State of Washington's Health Technology Assessment	http://www.hta.hca.wa.gov/
Evaluation of Genomic Applications in Practice and Prevention	http://www.egappreviews.org/
Cochrane Collaboration	http://www.cochrane.org/
National Institute of Health and Clinical Excellence	http://www.nice.org.uk/

Details regarding the significance of each agency and institution are set forth in greater detail, below. As peer-reviewed articles that specify the methodology for the evaluation of technologies used to screen diseases are difficult to find, these websites set forth the current criteria and standards for evaluating relevant healthcare technologies that are targeted toward screening for particular diseases.

- **Centers for Medicare and Medicaid Services (CMS) –** CMS is the largest public payer in the United States. As such, it has a great deal of influence over which new healthcare technologies will be covered and paid for by the Medicare Program. As many chronic diseases affect individuals that are 65 years or older, HTA decisions made by CMS are important in determining which technologies will be deemed acceptable for integration into clinical management of the Medicare patient population. Screening technologies may help to identify disease earlier in this segment of the patient population. As such, innovation in this area may alter the course of treatment for many diseases that affect Medicare patients. Medicare covers and pays for the care and treatment of medical conditions. One key issue is that Medicare does not cover and pay for the determination of risk factors or care for pre-conditions.
- **Agency for Healthcare Research and Quality (AHRQ) –** Health technology assessments are conducted for CMS by AHRQ. These HTA are used by CMS to help inform national coverage decisions for the Medicare program, as well as provide information to Medicare carriers. AHRQs HTA program uses current methodologies for evaluating the clinical utility of new technologies. HTAs are

based on a systematic review of literature, as well as qualitative and quantitative methods of synthesizing data from multiple studies. HTAs are conducted by AHRQ staff and can also be conducted in conjunction with one of the Evidence-based Practice Centers (EPCs).[17] In addition to its work for CMS, AHRQ has developed a program that contracts with institutions to participate in their EPC Program. The EPC's "review all relevant scientific literature on clinical, behavioral, and organization and financing topics to produce evidence reports and technology assessments. These reports are used for informing and developing coverage decisions, quality measures, educational materials and tools, guidelines, and research agendas."[18] When available, HTA topics are linked to corresponding information on the CMS website.

- **Blue Cross Blue Shield Association Technology Evaluation Center (BCBSATEC)** – Blue Cross Blue Shield is one of the largest private payers in the United States. BCBSATEC is one of the fourteen EPCs that are currently funded by AHRQ.[19]
- **U.S. Preventative Services Task Force (USPSTF) –** Since 1998, AHRQ has sponsored the USPSTF. USPSTF was first convened by the U.S. Public Health Service in 1984, and is the leading independent panel of private-sector experts in prevention and primary care. The USPSTF conducts rigorous, impartial assessments of the scientific evidence for the effectiveness of a broad range of

[17] Agency for Healthcare Research and Quality, Evidence-based Practice Program: Synthesizing scientific evidence to improve quality and effectiveness in healthcare. http://www.ahrq.gov/clinic/epc/. Accessed August 16, 2009.

[18] *Ibid*. http://www.ahrq.gov/clinic/epc/. Accessed March 15, 2009.

[19] *Ibid.* http://www.ahrq.gov/clinic/epc/epcenters.htm Accessed March 15, 2009.

clinical preventive services, including screening, counseling, and preventive medications. Its recommendations are considered the "gold standard" for clinical preventive services.[20] The mission of the USPSTF is to evaluate the benefits of individual services based on age, gender, and risk factors for disease; make recommendations about which preventive services should be incorporated routinely into primary medical care and for which populations; and identify a research agenda for clinical preventive care.[21]

- **National Library of Medicine (NLM) and National Institutes of Health (NIH) –** These agencies have information that discusses the foundation of HTA and other ancillary policy topics that are relevant to the field of HTA, such as comparative effectiveness research (CER), in the United States.
- **State of Washington's Health Technology Assessment (WAHTA) –** This agency is included because it provides an example of a state's process for introducing healthcare technology assessment to its population.
- **California Technology Assessment Forum (CTAF)** – "A community forum dedicated to objectivity and transparency regarding medical technologies; a community forum for dialog and decisions regarding the safety and effectiveness of new and emerging technologies."[22]
- **Evaluation of Genomic Applications in Practice and Prevention (EGAPP)** – EGAPP is "an initiative launched in 2004 to support a coordinated, systematic process for evaluating genetic tests and other genomic applications that are in

[20] "About USPSTF". http://www.ahrq.gov/clinic/uspstfab.htm. Accessed: August 2, 2009.

[21] *Ibid.* Accessed: August 2, 2009.

[22] CTAF website: http://www.ctaf.org/. Accessed: February 27, 2010.

transition from research to clinical and public health practice in the United States. The EGAPP Working Group was established in 2005 to support the development of a systematic process for assessing the available evidence regarding the validity and utility of rapidly emerging genetic tests for clinical practice. This independent, multidisciplinary panel prioritizes and selects tests, reviews CDC-commissioned evidence reports and other contextual factors, highlights critical knowledge gaps, and provides guidance on appropriate use of genetic tests in specific clinical scenarios."[23]

- **The Cochrane Collaboration –** The Cochrane Collaboration is dedicated to improving global healthcare decision-making by conducting systematic reviews of the effects of specific healthcare interventions. The results of these reviews are published in The Cochrane Library.[24]
- **National Institute of Health and Clinical Excellence (NICE) –** NICE is the technology assessment agency in the United Kingdom. It is included to provide insight as to how HTA is conducted in a national health setting outside of the United States. NICE is unique from United States technology assessment groups in that NICE reviews technologies prospectively for the United Kingdom before the National Health Service employs the particular technology. In the United States a technologies are typically retrospectively reviewed by technology assessment agencies after the technology has been marketed and is in use.

[23] Evaluation of Genomic Applications in Practice and Prevention. http://www.egappreviews.org/. Accessed: August 2, 2009.

[24] The Cochrane Collaboration. http://www.cochrane.org/. Accessed: August 2, 2009.

Literature Selection

Information researched included the types of criteria HTA organizations apply to technologies used to screening for disease. The reviewed articles encompass specific types of technology assessment criteria used to evaluate diagnostic and screening products, are not published prior to 2000; and are published in peer-reviewed journals. Diagnostic products were included to capture any references that might include information about both diagnostic and screening products, as well as guard against possible mischaracterization of technology. Articles were only included if they were peer-reviewed. In addition, the HTA data searched on organization websites had a similar restriction of not published or posted prior to 2000 in order to restrict the search to newer technologies used for disease screening. This date restriction was used because these technologies are relatively new and it was used to restrict the size of the retrieved material. Published and posted articles and reviews available HTA for relevant technologies older than 2000 were only used as historical sources.[25]

Exclusion criteria were articles that do not discuss healthcare technologies used to screen for diseases or conditions, not peer-reviewed, in foreign languages, and older than 2000. A review of citations, and then abstracts, provided the process for the initial selection of articles, with the goal of reviewing the full text of 15-20 articles to determine inclusion. On the websites, assessments that were published before 2000 were excluded from review. In addition, the specific research criteria for the website search was limited

[25] According to AHRQ's database of technology assessments, technology assessments conducted prior to 2004 may be outdated due to more recent research findings not included in the assessments and should be viewed cautiously in current medical practice. They are maintained by AHRQ for archival purposes only. I have selected an article inclusion date limit of 2000 to evaluate whether or not a history can be established regarding how the technology assessment criteria for diagnostic and screening products may have evolved for more current assessments. http://www.ahrq.gov/clinic/techix.htm Accessed February 22, 2009.

to criteria used to conduct HTA for screening products. Technology assessment agencies require published studies for review; however, some agencies also review material that is not peer-reviewed (e.g., summaries of evidence prepared by research associates employed by the agency, thought-leader presentations at meetings, etc). These types of additional materials are excluded from this literature review.

The keyword searches yielded a list of citations. These citations were reviewed for relevance to the field of HTA, as well as HTA involving screening tests. If a citation appeared relevant, the abstract was reviewed for relevance. If the abstract was deemed relevant, the full text of the article was retrieved and reviewed.

If the article provided information regarding the HTA criteria, the results of the HTA were reviewed to derive ideas regarding the validity or appropriateness of the methods. This informed the research objective and may serve as a baseline for comparing different HTA organizations, if relevant, to establish best practices for assessing technologies developed to screen for disease during interviews.

Literature Review Results

This section describes the overall search results for this literature review and corresponding themes of research. The search results were generated using the key word searches set forth in Tables 1 and 2, above, and was conducted using PubMed and Google Scholar. In addition, search results were also compiled from the websites of agencies and institutions set forth in Table 3, above.

Description of Search Results

Based on the results of the preliminary key word searches set forth in Table 2, the key words were grouped and divided into three search cohorts. Key word searches were

repeated employing date inclusion criteria to narrow the search for relevant articles published between: January 1, 2000 and January 1, 2009. In addition, based on the depth and breadth of the search results, the cohort was either deemed excluded from further evaluation (e.g., Cohort 1), or included for further evaluation (e.g., Cohorts 2 and 3). The results of the key word searches from the three search cohorts and whether they were included or excluded from evaluation are set forth below in Table 4.

Table 4. Search Cohort Inclusion & Exclusion for Word Search & Results

Search Cohort Inclusion / Exclusion	Key Words	Search Results
Cohort 1 Excluded from Evaluation	Healthcare Technology Assessment Criteria	767
	Healthcare Technology Evaluation Criteria	1,079
	Healthcare Technology Assessment Diagnostic	1,247
	Healthcare Technology Evaluation Screening	681
Cohort 2 Included in Evaluation	Healthcare Technology Assessment Diagnostic Test	26
	Healthcare Technology Assessment Screening Test	155
Cohort 3 Included in Evaluation	Healthcare Technology Evaluation Diagnostic Test	36
	Healthcare Technology Evaluation Screening Test	246

The key word search terms were the most broad in Cohort 1. The search focused on the type of criteria used in HTA and evaluations and was not limited by a type of condition or whether a diagnostic or screening product was used. Thus, although these two searches yielded a tremendous number of articles, 767 and 1,079, respectively, the topics of the articles included every type of healthcare technology assessment and evaluation conducted. As such, these searches were deemed too broad and were excluded from further evaluation.

By substituting the words “diagnostic” and “screening” for “criteria” in Cohort 1, yielded somewhat narrower results; however, still too broad for specific research. The word “diagnostic”, and variants thereof, are frequently used to indicate that a disease was diagnosed, not that a diagnostic test was used. Similarly, in the case of “screening”, this word, and variants thereof, are frequently used to explain that an x-ray or other form of

screening technology, and not test used to screen for disease, was used. Because the results from these four key word searches were not limited to technology assessments and evaluations involving diagnostic or screening test products, they were excluded from further evaluation.

Key word limitations were then applied to narrow the search results to those HTA that involved either a "healthcare technology assessment" with a "diagnostic test" or "screening test" (Cohort 2) or a "healthcare technology evaluation with a "diagnostic test" or "screening test" (Cohort 3) published between the dates: January 1, 2000 and January 1, 2009. While these key word searches yielded a significant number of articles, they were limited to the research topic. Consequently, article titles and, if relevant, abstracts, were reviewed for the key word search terms included in Cohort 2 and 3.

A total of 463 results were reviewed by citation and title. If the citation and/or title were reflective of the research question, the abstract was examined to determine if it was relevant. Of the 463 articles, a total of 17 articles were deemed relevant for further review. These 17 articles were retrieved and reviewed in their entirety. Upon reviewing these 17 articles, 12 were deemed relevant for the purposes of this literature review.

Table 5. Search Cohort Inclusion for PubMed Key Word Search & Results

Search Cohort Inclusion / Exclusion	Key Words	Search Results	Relevant Articles
Cohort 2 Included in Evaluation	Healthcare Technology Assessment Diagnostic Test	26	7
	Healthcare Technology Evaluation Screening Test	155	1*
Cohort 3 Included in Evaluation	Healthcare Technology Assessment Diagnostic Test	36	8
	Healthcare Technology Evaluation Screening Test	246	2
Total Articles		**463**	**17**

** Indicates repetitive citations, therefore excluded from Total Articles calculation.*

Description of Search Results – Website

Web searches were conducted for the agencies and institutions set forth in Table 3. The websites for these agencies and institutions were chosen due to their influence and

participation in HTA in the United States. The Centers for Medicare and Medicaid Services (CMS) represents the largest public payer in the United States for people age of 65 and older. CMS' technology assessments are conducted by the Agency for Healthcare Research and Quality (AHRQ). As a result, the websites for both of these institutions are included for review due to the breadth of the patient population that may be affected by HTA for diagnostic and screening products.

Using the key word searches from Cohort 2 and Cohort 3, the majority of the articles retrieved in this search outlined specific technology assessments for particular products. The specific technology assessment criteria for evaluating diagnostic and screening were not specified. Thus, when specific products are selected to research the technology assessment process, the search will be easily narrowed to the specific product and review articles that are relevant to the particular disease state and technology reviewed.

Similarly, AHRQ, BCBSTEC, USPSTF, EGAPP, DERP, State of Washington's Health Technology Assessment, CTAF, and the United Kingdom's NICE websites also provided information regarding technology assessments that had been completed, or are in process, for a wide variety of technologies. Thus, once a product is selected, these resources will also be helpful in identifying the means by which technologies for specific diseases have been evaluated.

Peer-Reviewed Journal Themes

A total of 12 peer-reviewed journal articles were identified that appeared relevant to the research question. None of the articles were directly on point; however, they did either address a component of technology assessment, complexity of evaluating

diagnostic and screening technologies, cost-effectiveness and how it might relate to technology assessment, or an HTA for a specific disease or condition.

Among the 12 articles, there were 4 prevalent themes: (1) value and attainability of outcomes evidence; (2) evaluation of cost-effectiveness of new technologies; (3) evaluation of patient perspectives; and (4) danger of developing fragmented healthcare policy. Each of these themes is discussed in detail, below.

(1) Value of Outcomes Evidence.

One of the prevalent themes in the literature involves the type of evidence that should be collected and evaluated in order to assess a screening technology. Many authors recognize the traditional strength of outcomes evidence collected during a randomized clinical trial; however, the cost involved in collecting this type of evidence for screening technologies is prohibitive.[26] Another data and cost consideration is how long to collect evidence for a screening technology. For example, should data be collected until the disease presents itself, or should data collection continue through therapeutic treatment until the eventual death of the patient? Thus, the key issue is distinguishing the ideal: until natural death of all cohort members versus a feasible study (i.e., IRB-approved study that is affordable to conduct).

Expert opinion regarding the level of evidence needed for a screening technology was not found. However, the literature did provide guidance relative to diagnostics. Bruns and Hartmann set forth the complexity of determining an appropriate level of evidence in evaluating diagnostic products. While outcomes studies have worked well for therapeutic interventions, the study of laboratory-related outcomes is more complex

[26] Bruns, D. E. (2001). Laboratory-related outcomes in healthcare. *Clinical Chemistry, 47*(8), page 1548.

because multiple steps occur between the time of the test and the eventual healthcare outcome.[27] Outcomes studies address questions such as: "is use of test X associated with outcome Y?"[28] As a result, a question might be whether a test occurring early on in the treatment cycle decreases eventual length of stay. Bruns argues that outcomes studies, such as randomized controlled trials, may not be appropriate measures for the effectiveness of tests because they are cost prohibitive. As a result, while outcomes studies might be used in evaluating these types of technologies, they may not always be possible.[29] Hartmann concludes the opposite, that "prospective measurement of outcomes is essential to make clinical decisions about test methods and to guide policy.[30]

For 30 years, Evidence-based Practice Centers (EPCs) have used a 6-level framework to evaluate diagnostic technologies. This framework consists of the following levels: (1) technical feasibility and optimization; (2) diagnostic accuracy; (3) diagnostic thinking impact; (4) therapeutic choice impact; (5) patient outcome impact; and (6) societal impact.[31] Sunshine summarizes applies the 6-level test to radiological diagnostic imaging technologies in Table 6, below.

---

[27] Bruns, D. E. (2001). Laboratory-related outcomes in healthcare. *Clinical Chemistry, 47*(8), 1547-1552.

[28] *Ibid*, page 1548.

[29] *Ibid*, page 1552.

[30] Hartmann, K. E., Nanda, K., Hall, S., & Myers, E. (2001). Technologic advances for evaluation of cervical cytology: Is newer better? *Obstetrical & Gynecological Survey, 56*(12), 765-774.

[31] *Ibid*, page 1049 and 1054.

Table 6. Hierarchy of Efficacy for Diagnostic Technologies

EPC Level	Typical Measures
1. Technical Efficacy	• Resolution of line pairs • Pixels per millimeter • Section thickness • Noise level
2. Diagnostic Accuracy	• Sensitivity • Specificity • Area under the receiver operating characteristic curve
3. Diagnosis	• Percentage of cases in which image is judged helpful in making the diagnosis • Percentage of cases in which diagnosis made without the test is altered – or altered substantially – when information from the test is received
4. Treatment	• Percentage of cases in which image is judged helpful in planning patient treatment • Percentage of cases in which treatment planned without the test is changed after information from the test is received
5. Patient Health Outcomes	• Percentage of patients improved with test conducted compared with that improved without test conducted • Percentage difference in specific morbidities with test compared with those without • Mean increase in quality-adjusted life years with test compared with that without
6. Societal Value	• Cost-effectiveness from a societal perspective • Cost per life saved, calculated from a societal perspective

Sunshine sets forth that the basic questions of the efficacy of a diagnostic intervention are, "[h]ow much does this do to improve the health of people?" and "How much does it cost for that gain in health?"[32] From this perspective, one way to consider the gap between diagnostic accuracy and outcomes is to use the 6-level "hierarchy of efficacy" test.[33] Unfortunately, there is rarely sufficient study evidence to satisfy each of

[32] Sunshine, J. H., & Applegate, K. E. (2004). Technology assessment for radiologists. *Radiology, 230*(2), 309-314.

[33] *Ibid*, page 309. Citing the 6-level test belonging to: Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991; 11:88-94. Also citing: Thornbury JR. Clinical efficacy of diagnostic imaging: love it or leave it. AJR Am J Roentgenol 1994; 162:1-8.

the levels. Sunshine describes decision analysis and cost-effectiveness analysis as two other ways to measure the efficacy of diagnostic technology.[34]

(2) Evaluation of Cost-Effectiveness of New Technologies.

A second prevalent theme in the peer-reviewed literature is that screening technologies may identify the presence of disease earlier. As a result, physicians and patients may have more treatment options available for the treatment of the disease at an earlier phase of disease development. Consequently, this may lead to savings in the overall treatment of the disease. A counter-argument exists that earlier discovery of disease results in higher cost because although the disease is identified earlier on in the development of the disease, it ultimately results in requiring the disease to be treated for a longer duration of time, thus incurring higher overall cost.[35]

Tatsioni argues that the main purpose of a test is to help guide patient management decisions in order to improve patient outcomes. As a result, because tests are important to healthcare decision-making, they should be evaluated with the same amount of effort as a therapeutic intervention.[36] Supporters of new technologies frequently indicate that the use of these technologies would improve outcomes, avoid unnecessary procedures and possibly reduce costs.[37] However, these positions depend on early screening for a condition for which there is an effective intervention to either cure or slow the progression of symptoms in a substantial subset of patients who can be

---

[34] Sunshine, page 310.

[35] Tatsioni, A., Zarin, D. A., Aronson, N., Samson, D. J., Flamm, C. R., Schmid, C., et al. (2005). Challenges in systematic reviews of diagnostic technologies. *Annals of Internal Medicine, 142*(12 Pt 2), 1048-1055.

[36] *Ibid*, page 1048.

[37] *Ibid*, page 1053.

identified prior to screening. Further, the disease or condition should be sufficiently serious to warrant a medical intervention. For an example, the stated terms of payer's plan might specifically indicate the scenarios in which an intervention is warranted or not.

Bodenheimer focuses on describing methods to control costs while ensuring quality care.[38] The purpose of this research study is not to reflect on process measures and patient experience. As a result, a discussion about quality of care is outside the range of the evaluation of this research study. With this taken under consideration, while this article does not specifically address screening technologies, it does outline several strategies to contain healthcare costs by employing various disease management approaches, as well as outlining the importance of improved technology assessment to reduce the overuse of new technologies that may not be appropriate. Specifically, Bodenheimer sets forth six interrelated strategies to manage cost and increase quality: (1) programs that target the percent of the overall population that incur significant healthcare costs; (2) disease management programs used to prevent expensive complications of chronic conditions; (3) ways to prevent medical errors; (4) increasing the strength of primary care; (5) decision-support tools to help avoid services that are not needed; and (6) better diffusion of technology assessment.[39]

Based on an analysis of Medicare data, Bodenheimer indicates 10% of the population is responsible for 70% of healthcare expenditures.[40] This high-cost 10% of the

[38] Bodenheimer, T et al. High and rising health care costs. Part 4: can costs be controlled while preserving quality? Ann.Intern.Med., 2005, 143, 1, 26-31.

[39] *Ibid.*

[40] *Ibid.*

Medicare patient population typically have long term chronic diseases that could have been mitigated in terms of better quality care and decreased healthcare expenditure if the patient had access to primary care. However, the purview of the Medicare program is to provide treatments to patient, not to deliver preventative care via screening and diagnosis of disease. Bodenheimer argues that the use of these six strategies can play a significant role in decreasing overall healthcare cost by reducing hospital use and increasing quality of care to patients on an ongoing basis.[41]

Thus, while this article does not exclusively focus on technology assessment or technology assessment for screening products, it does set forth the premise that technology assessment should be used to help direct appropriate technology to appropriate patients and postulates that this will result in decreased healthcare expenditures. This theme is congruent with the goal of screening products to identify disease earlier to better understand and select appropriate treatment options for relevant patient populations. While this conclusion might be true for therapeutic interventions, Rogowski sets forth that favorable cost-effectiveness ratios have not yet been obvious for genetic testing technologies.[42]

(3) Evaluation of Patient Perspectives.

A third theme that appeared in the peer-reviewed literature involved including the patient's perspective in evaluating the overall effectiveness of a technology. For example, although a technology such as virtual colonoscopy might be extremely effective in screening for colorectal cancer. But, if patients are not adherent in getting the

---

[41] *Ibid.*

[42] Rogowski, W. (2007). Current impact of gene technology on healthcare. A map of economic assessments. *Health Policy (Amsterdam, Netherlands), 80*(2), 340-357.

procedure, then the offering screening program is not efficacious because patients are not receiving the test.[43] As a result, patient perspectives might be a valuable criterion to consider in evaluating the effectiveness of new procedures. Patient perspectives can relate to satisfaction and quality of life. These factors influence adherence and patient outcomes.

For screening tests, Schneider, Birch, Inadomi, and Brawley take into account the perspective of the patient in selecting an intervention. Schneider tries to determine the actual usefulness of high-cost imaging technologies in Germany by evaluating the level of specificity of a patient's complaint. Birch discusses taking into account patient preference in evaluating whether or not to pursue a particular intervention as it may affect the patient's ultimate well-being.[44, 45] Inadomi examines how patient adherence to going to get a screening test might affect the use of a test that is highly effective in identifying disease early.[46] Lastly, Brawley sets forth that a screening test should only be used when the potential for benefit outweighs the potential for harm.[47] Specifically, Brawley sets forth that "[t]o be of benefit, screening must find disease earlier and lead to an efficacious treatment, and earlier use of the efficacious treatment must offer better outcome compared to treatment at the onset of symptoms. The benefits must also outweigh the

---

[43] Inadomi, J. M. (2008). Taishotoyama symposium barriers to colorectal cancer screening: Economics, capacity and adherence. *Journal of Gastroenterology and Hepatology, 23 Suppl 2*, S198-204.

[44] Schneider, A., Rosemann, T., Wensing, M., & Szecsenyi, J. (2005). Physicians perceived usefulness of high-cost diagnostic imaging studies: Results of a referral study in a german medical quality network. *BMC Family Practice, 6*(1), 22.

[45] Birch, S., & Ismail, A. I. (2002). Patient preferences and the measurement of utilities in the evaluation of dental technologies. *Journal of Dental Research, 81*(7), 446-450.

[46] Inadomi, page 198.

[47] Brawley, O. W., & Kramer, B. S. (2005). Cancer screening in theory and in practice. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 23*(2), 293-300.

harms."[48] Brawley cites that there are examples in which screening tests and treatment found cancer at an earlier stage; however, while survival from the date of diagnosis was longer, mortality was not affected and a net harm was incurred.[49]

Schneider determined that diagnostic technologies were most effective in situations in which the patient had a specific complaint. The usefulness of the technology decreased if the patient had a vague complaint for diagnosis. As a result, he concludes: "quality improvement should focus on patients with unexplained complaints to avoid expensive, unnecessary or dangerous diagnostic investigations."[50] The data from this study supports an investigation into whether specificity of the patient's complaint should be factored into the level of evidence used when evaluating a diagnostic tool. Inadomi concludes that if a patient is not adherent to a highly effective screening test, we should question the overall effectiveness of a test that patients do not want to take.[51]

(4) Danger of Developing Fragmented Healthcare Policy.

The fourth theme identified in the literature was a risk of developing fragmented healthcare policy in conducting HTA of screening technologies. Because there are many different agencies conducting HTA, these agencies have different missions, visions, and funding. Further, organizations that use HTA as a means for making decisions about which technologies to offer to beneficiaries employ different strategies for applying HTA. These organizations have different objectives and responsibilities in the type of healthcare that is owed to their beneficiaries. As a result, if access to healthcare for all

---

[48] *Ibid*, page 293.

[49] *Ibid.*

[50] Schneider, page 5.

[51] Inadomi, page 203.

patients should be equal, currently organizations that provide healthcare to beneficiaries do not have a point of congruence in the level of evidence required for new technologies, there is a substantial risk of creating policy that is not stable. Conversely, if it is acceptable to have different types of healthcare offerings available to beneficiaries, then it is acceptable to have different policies. The overarching danger results when fragmented healthcare policy results in stymied innovation and lack of reasonable access to care.

Vermeulen sets forth the overall policy theme regarding the danger of fragmentation in the HTA policy environment. Specifically, Vermeulen discusses the challenges faced when involving different authoritative bodies with different agendas in evaluating new technologies. Although this paper refers to agencies and technology assessment policy in Belgium, the results are translatable to the United States. The conclusion that involving multiple authorities with different agendas results in fragmented policy is a solid conclusion that is a warning sign for the need to develop congruencies between HTA agencies.[52]

Agency and Institution Website Themes

In the United States, there are several agencies that conduct HTAs. AHRQ conducts HTAs for CMS; in addition, they work in collaboration with private institutions to conduct HTAs through their Evidence-based Practice Centers (EPC) Program. In addition, there are several private institutions that conduct HTAs. The common theme throughout the evaluation of these institution websites is that there is congruence between

[52] Vermeulen, V., Coppens, K., & Kesteloot, K. (2001). Impact of health technology assessment on preventive screening in belgium: Case studies of mammography in breast cancer, PSA screening in prostate cancer, and ultrasound in normal pregnancy. *International Journal of Technology Assessment in Health Care, 17*(3), 316-328.

institutions that HTAs are important in evaluating the contribution that a particular technology may make. In addition, there is congruence for the criteria used to evaluate therapeutic interventions. However, one area that is developing is the review of screening technologies. An overview of each institutional website as it relates to HTA is set forth below.

The website for the Centers for Medicare and Medicaid Services (CMS) does not have a direct link to either "technology assessment" or "technology assessment for diagnosis or screening". The results of both of these searches provide links to the Agency for Healthcare Research and Quality (AHRQ) that reviews specific technologies for CMS. These HTAs are used by CMS to formulate national coverage decisions for the Medicare patient population, as well as to provide information to Medicare carriers. The CMS website does not contain content that specifies the general guidelines or assessment criteria for products used to diagnose or screen for disease.

The AHRQ website contains information about the Center for Practice and Technology Assessment (CPTA).[53] Specifically, CPTA "helps to narrow the gap between what is known from research about effective and efficient clinical care and what is practiced in health care settings."[54] The organization "was established to serve as a single contact for organizations and individuals searching for comprehensive evidence reviews on health conditions, treatments and technologies."[55] CPTA has four program activities: (1) Evidence-based Practice Centers (EPCs) that conduct technology assessments; (2) National Guideline Clearinghouse that is a source of clinical practice

[53] http://www.ahrq.gov/about/cpta/cptafact.htm Accessed: April 12, 2009.

[54] *Ibid,* Accessed: April 12, 2009.

[55] *Ibid,* Accessed: April 12, 2009.

guidelines; (3) US Preventative Services Task Force (USPSTF) which educates primary care physicians, health plans, and payers about what services should be included in periodic health examinations; and (4) Research and Evaluation focused on methodologies used in conducting systematic, evidence-based reviews and syntheses, as well as approaches used for implementing evidence-based clinical information into the health care delivery system.[56]

The EPCs and USPSTF are the most relevant of AHRQ's four programs to understand the technology assessment criteria that might be used for new screening technologies. The function of the EPCs is to "develop evidence reports and technology assessments about clinical topics that are common, expensive, and/or significant for the Medicare and Medicaid populations." Currently, there are 14 EPCs that are contracted in the United States and Canada to participate in the program for 5-year intervals. Table 7 sets forth the current EPCs that are under contract with AHRQ.

Table 7. AHRQs Evidence-Based Practice Centers

Blue Cross and Blue Shield Association, Technology Evaluation Center
Duke University*
ECRI Institute*
Johns Hopkins University
McMaster University
Minnesota Evidence-based Practice Center
Oregon Evidence-based Practice Center**
RTI International – University of North Carolina
Southern California
Tufts – New England Medical Center*
University of Alberta*
University of Connecticut
University of Ottowa
Vanderbilt University

* EPCs that focus on HTAs for CMS.
** EPCs that focus on evidence reports for USPSTF.

Evidence reports and technology assessments are based on rigorous review of all relevant scientific literature and include specific analyses, such as meta-analyses and cost

[56] *Ibid,* Accessed: April 12, 2009.

analyses, as warranted. Use of these reports and assessments will help reduce inappropriate variations in medical care and improve the overall quality and efficiency of the health care system.[57] The EPC section of the website provides specific HTAs for products. Thus, once a screening intervention is selected for comparison across institutions, a search of this section of the website will be helpful in identifying HTAs that have been performed relevant to diagnosing or screening for that disease. These specific assessments might provide insight as to whether HTA criteria are modified in reviewing products used to screen for disease. However, the website does not contain content that specifies the assessment criteria for products used for screening. Rather, AHRQ and EPCs engage in systematic reviews of technologies.

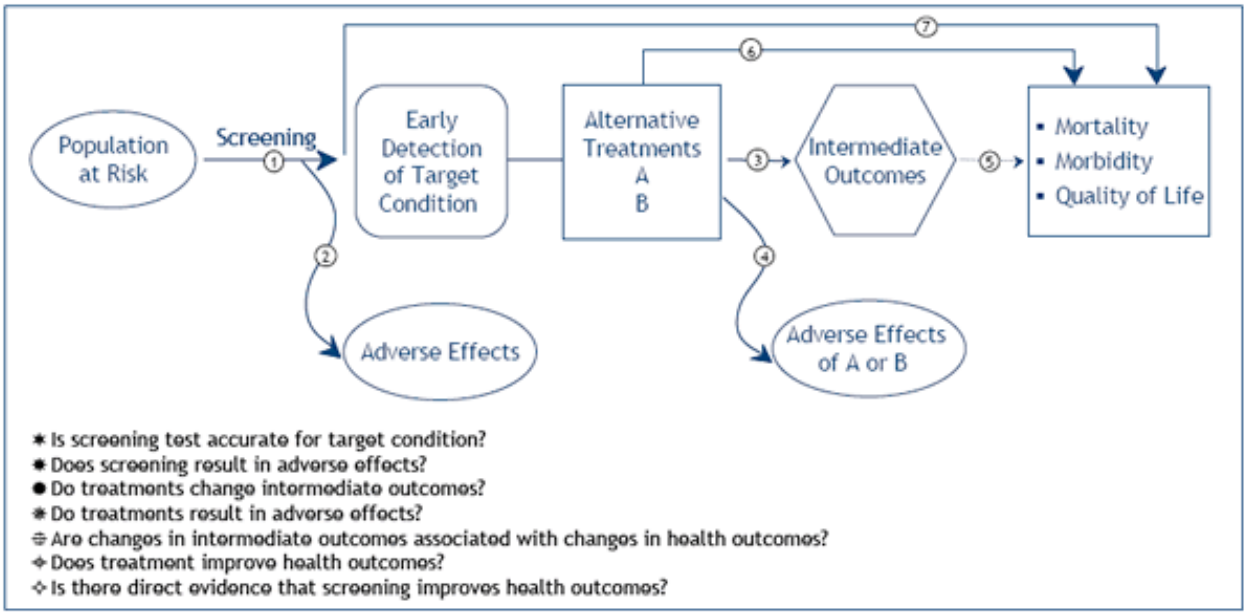
One portion of the website provided guidance in determining whether or not a technology should be reviewed. As applied to screening technologies, one component of the decision-making process in identifying technologies to review is the development of a causal pathway or analytic framework.[58] Developing the causal pathway or analytic framework is a means for specifying evidence questions for new technologies. Specifically, a graphical display can show direct and indirect links between interventions and outcomes and are particularly useful for screening interventions. Because screening interventions involve a series of decisions that could each be a topic of an evidence question, the graphic is helpful in illustrating where questions of evidence might occur. A sample causal pathway is set forth in Figure 1, below.

---

[57] *Ibid,* Accessed: April 12, 2009.

[58] http://www.ahrq.gov/clinic/epcpartner/epcpartner2.htm#fig1 Accessed: August 15, 2009.

Figure 1. General Causal Pathway – Screening Procedure and Alternative Treatments[59]



This figure provides some guidance as to the possible evaluation points of a screening technology.

The other section of AHRQs website that is useful is the USPSTF. The USPSTF "is an independent panel of experts in primary care and prevention that systematically reviews the evidence of effectiveness and develops recommendations for clinical preventive services."[60] The USPSTF portion of AHRQs website sets forth tools and resources that clinicians can integrate into their practice in order to develop best practices for prevention. Once screening interventions are selected for review, this website will be searched for USPSTFs review of the particular intervention.

The Blue Cross Blue Shield Association's Technology Evaluation Center (BCBSATEC) website provides a record of the technology assessments conducted by the

[59] http://www.ahrq.gov/clinic/epcpartner/epcpartner2.htm Accessed: August 15, 2009. Citing: Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force. A review of the process. *Am J Prev Med* 2001;20(3S):21-35.

[60] http://www.ahrq.gov/clinic/prevenix.htm Accessed: August 15, 2009.

BCBSATEC. These reports are designed to help readers understand the scientific evidence on the effectiveness of different treatments and tests, as well as make informed decisions about healthcare.[61] The five BSBSATEC criteria are set forth on the website: (1) the technology must have final approval from the appropriate governmental regulatory bodies; (2) the scientific evidence must permit conclusions concerning the effect of the technology on health outcomes; (3) the technology must improve the net health outcome; (4) the technology must be as beneficial as any established alternatives; and (5) the improvement must be attainable outside the investigational settings.[62] Like the AHRQ website, there is a library of drugs, medical devices, and biotechnology products that have received a technology assessment; however, the website does not contain content that specifies the assessment specific criteria for products used to screen for disease.

Like the BCBSATEC website, the National Library of Medicine (NLM) website has information regarding HTAs. While the NLM does not conduct technology assessments like AHRQ or BCBSATEC, it does have a resource that describes the history of health technology assessment by Cliff Goodman, PhD, a noted expert in the HTA field. In order to learn more about resources about HTA and whether the criteria for the assessment might change for technologies used to screen for disease, I contacted Dr. Goodman directly.[63] Dr. Goodman has offered to help identify appropriate products to serve as the focus for additional research and subsequent technology assessments that

---

[61] http://www.bcbs.com/blueresources/tec/ Accessed: April 12, 2009.

[62] http://www.bcbs.com/blueresources/tec/tec-criteria.html Accessed: April 12, 2009.

[63] Goodman, C. HTA 101 Introduction to Health Technology Assessment. 2004. http://www.nlm.nih.gov/nichsr/hta101/ta10103.html#Heading4 Accessed: April 12, 2009.

might be helpful for the research question.[64] As a preliminary suggestion, Dr. Goodman recommended examining technology regarding computed tomographic colonography, a type of "virtual colonoscopy" technology.

The State of Washington's Health Technology Assessment (WAHTA) website provides detail regarding the state's program. "The primary purpose of [WA]HTA is to ensure medical treatments and services paid for with state health care dollars are safe and proven to work."[65] Because of the cost of conducting technology assessments, only 10 have been conducted by WAHTA within the past two years. These assessments are posted on the website and some are still open for review and comment. The topics are set forth below in Table 8.

Table 8. State of Washington Health Technology Assessments

Technology	Category	Date	Status
Cardiac Stent	Therapeutic	5/8/09	Review Open
Computed Tomographic Angiography	Diagnostic	11/14/08	Review Open
Artificial Discs	Therapeutic	10/17/08	Review Open
Implantable Infusion Pumps	Therapeutic	8/15/08	Decision Complete
Arthroscopic Knee Surgery	Therapeutic	8/15/08	Decision Complete
Discography	Diagnostic	2/15/08	Decision Complete
Virtual Colonoscopy or Computed Tomographic Colonography	Diagnostic	2/15/08	Decision Complete
Lumbar Fusion	Therapeutic	11/16/07	Decision Complete
Pediatric Bariatric Surgery	Therapeutic	8/24/07	Decision Complete
Upright/Positional MRI	Diagnostic	5/18/07	Decision Complete

Six of these technologies address therapeutic products, 4 address diagnostic products.

California Technology Assessment Forum (CTAF) "is a public service forum that assesses new and emerging medical technology."[66] This organization believes that the growth of new technologies represents the most important change for the delivery of

[64] Teleconference with Dr. Cliff Goodman, April 13, 2009.

[65] http://www.hta.hca.wa.gov/about.html Accessed: April 12, 2009.

[66] http://www.ctaf.org/section/aboutus Accessed: February 27, 2010.

healthcare in America. While technologies have a positive effect for patients, there is also a risk of misuse and overuse of technology. CTAF is an active proponent of evidence-based medicine. CTAF is spearheaded by Blue Shield of California Foundation which manages the technology assessment reviews and organizes all CTAF meetings and events.[67] CTAF has assessed a significant number of technologies. Interestingly, like WAHTA, virtual colonoscopy was reviewed; however, CTAF distinguished this technology as a screening technology, while WAHTA identified it as a diagnostic. This indicates that there may be confusion of terminology between agencies assessing the same technology.

The primary criteria for these assessments include: potential patient harm/safety concerns; concerns about therapeutic efficacy or diagnostic accuracy and appropriateness of outcomes for patients; and estimated total direct cost per year (estimated increase/decrease).[68] The secondary criteria for these assessments include: number of persons affected per year, severity of condition treated by technology, policy related urgency/diffusion concern, potential or observed variation; and special populations/ethical concerns.[69] Because these criteria are broader than those of BCBSATEC, the technology assessments may provide guidance once a specific screening product is selected. However, the website does not contain content that specifies the assessment criteria for screening products.

The website for the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) provides guidance on the use of genetic tests in specific clinical

---

[67] *Ibid.*

[68] http://www.hta.hca.wa.gov/documents/prioritization_criteria.pdf Accessed: April 12, 2009.

[69] *Ibid.*

settings.[70] There is a recent report posted regarding Outcomes of Genetic Testing in Adults with a History of Venous Thromboembolism.[71] Once a specific screening procedure is selected for review, this website will be reviewed for any relevant reports.

The Cochrane Collaboration seeks to improve global healthcare by conducting systematic reviews of technology.[72] The website will be searched once relevant screening products are selected for review and comparison across organizations that have conducted HTAs.

Like AHRQ, BCBSATEC, WAHTA, and CTAF, the National Institute of Health and Clinical Evaluation (NICE) website also sets forth their criteria for health technology assessments and a library of completed assessments. Guidance is provided in five key areas: cancer service, clinical guidelines, interventional procedures, public health, and technology appraisals. Of these five areas, "technology appraisals" specifies those technologies that have published reports, as well as those that are slated for review. As with the other websites, there is no content that provides detail about the assessment criteria for products used to screen for disease.

In each of the websites, the general technology assessment criteria were set forth; however, the bulk of the published material involves specific technology assessments for particular technologies. As a result, a selection of technologies directed toward a particular product used to treat a specific disease is recommended in order to compare the evaluation across agencies for the same technology used to screen for disease.

---

[70] http://www.egappreviews.org/ Accessed: August 15, 2009.

[71] *Ibid.*

[72] http://www.cochrane.org/ Accessed: August 15, 2009.

Role of Comparative Effectiveness Research in HTA

Comparative Effectiveness Research (CER) emerged from the concern that when multiple treatment options exist, there is limited evidence of the effectiveness of one particular product over another when used in a routine care setting. In this context, "effectiveness" is defined as "whether an intervention does more good than harm when provided under usual circumstances of health care practice."[73] CER is receiving a significant amount of attention due to provisions in the American Recovery and Reinvestment Act of 2009 (ARRA).[74] ARRA provides approximately $1.1 billion toward conducting and supporting CER, including establishing a Federal Coordinating Council on Comparative Effectiveness Research (FCCCER).[75]

Both physicians and health care payers have an interest in CER. Physicians generally have choices among products to prescribe or use to treat a particular condition. Therefore, when multiple treatment options are available, the question becomes: which product or treatment should be chosen because it is the most effective?[76] Similarly, health plans are interested in CER for two reasons. First, health plans have a vested interest in the health of their beneficiaries. Second, health plans have an interest in their financial health. As such, health plans want to structure utilization mechanisms, such as formularies and co-payments, to motivate physicians and patients to use the equally

---

[73] Schneewiss, S. (2007). Developments in post-marketing comparative effectiveness research. *Clinical Pharmacology & Therapeutics*, 82(2), 143-156, page 143.

[74] American Recovery and Reinvestment Act of 2009 (February 17, 2009). http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_bills&docid=f:h1enr.pdf. Accessed: January 17, 2010. Congress passed ARRA on February 13, 2009, and President Obama signed it into law on February 17, 2009.

[75] *Ibid.*

[76] *Ibid.*

effective, but less costly, treatment alternative.[77] CER is a mechanism to compare multiple options that are available in the market, while HTA is a mechanism to examine the risks and benefits of one particular product.

Literature Review Discussion

Description of Current Knowledge

The objective of this literature review was to determine how researchers and policy makers assess technologies related to screening products as they relate to a particular disease. A search of relevant literature and websites revealed that it is not immediately apparent whether the criteria for healthcare technology assessments for products that screen for a particular disease are any different from assessments for therapeutic products.

As presented earlier, there are three major thematic findings that have been revealed through the literature review process. These findings are set forth below in Table 9:

Table 9. Literature Review Thematic Findings

Finding	Thematic Finding
1	The overall value of outcomes evidence to evaluate screening products in a HTA is unclear.
2	While screening products may identify elevated risk of disease earlier and expand treatment options, it is not clear that the overall cost of care is reduced. Therefore, it is unclear whether cost or cost effectiveness should be a criterion for evaluating these technologies as a part of an HTA.
3	Whether patient experience should be a criterion in the HTA process of evaluating a new screening product is not considered.
4	There is a danger of further fragmented healthcare policy without consistent guidance.

A discussion of each of these thematic findings follows.

[77] *Ibid*, at 144.

Finding 1. The overall value and attainability of outcomes evidence to evaluate screening products in a HTA is unclear.

Understanding the appropriate level of evidence to evaluate screening technologies is the first essential step of an HTA. It appears from the existing peer-reviewed literature that there is a lack of agreement regarding what type of evidence might best support the HTA evaluation and eventual use of a screening technology. While outcomes evidence that is collected by conducting a randomized controlled clinical trial certainly produces the highest level of evidence, it is unclear whether this high standard should be applied to technologies that detect the presence of disease, rather than treat the disease.

This particular finding is discussed in the Bruns, Hartmann and Sunshine articles. Bruns argues that studies that yield outcomes evidence, such as randomized controlled trials, may not be appropriate measures for the effectiveness of diagnostic tests because they are cost prohibitive. As a result, while outcomes studies might be used in evaluating these types of technologies, they may not always be possible.[78] Hartmann concludes the opposite, that "prospective measurement of outcomes is essential to make clinical decisions about test methods and to guide policy."[79] Meanwhile, Sunshine evaluates the 6-level test used by EPCs and suggests rather than using outcomes evidence to evaluate these technologies, perhaps decision analysis and cost-effectiveness are two more appropriate measures.[80]

---

[78] *Ibid*, page 1552.

[79] Hartmann, K. E., Nanda, K., Hall, S., & Myers, E. (2001). Technologic advances for evaluation of cervical cytology: Is newer better? *Obstetrical & Gynecological Survey, 56*(12), 765-774.

[80] Sunshine, page 309.s

In a review of the websites, a comparison of reviews of a select number of technologies across organizations may yield an understanding of how the assessment criteria for technologies used to screen for disease might be different than the criteria for therapeutic technologies. The current knowledge gleaned from the identified websites supports this first thematic finding. Once a particular product is selected for comparisons across HTA agencies, differences in levels of evidence might become apparent. If there are differences in evidence levels, an evaluation as to why the levels are different might yield new approaches, beyond the collection of outcomes evidence, which might support the HTA of a screening product.

> ***Finding 2. While screening products may identify elevated risk of disease earlier and expand treatment options, it is not clear that the overall cost of care is reduced. Therefore, it is unclear whether cost or cost–effectiveness should be a criterion for evaluating these technologies as a part of an HTA.***

The evaluation of cost has been an area of interest for HTA in the United States. Traditionally, cost and cost effectiveness have been excluded as a review factor. However, in cases in which it is unclear whether a health benefit is being delivered, an analysis of cost would be helpful to distinguish whether establishing the presence of disease earlier results in better health outcomes that are cost effective.

This second finding is discussed in various forms in the Bodenheimer, Tatsioni and Rogowski articles.[81] The thesis of the Bodenheimer article is that there are strategies that involve some form of technology assessment that can be employed to reduce the risk of inappropriate patient care at a high cost to the healthcare system. The arguments in support of this finding in the Bodenheimer article are compelling, particularly in the

[81] Bodenheimer, et al. Page 30.

demonstration that in an analysis of the Medicare population, 10% of the population is responsible for 70% of healthcare expenditures.[82]

Tatsioni argues that the main purpose of a diagnostic test is to help guide patient management decisions in order to improve patient outcomes. As a result, because diagnostic tests are important to healthcare decision-making, they should be evaluated with the same amount of effort as a therapeutic intervention and cost should play a role in the evaluation.[83] By way of a baseline, Rogowski sets forth that favorable cost-effectiveness ratios have not yet been obvious for genetic testing technologies.[84]

The current knowledge from the published literature supports the finding that evidence-based medicine and coordinated care can result in better quality care, less illness and lower healthcare expenditure. In each of the articles, technology assessment comes to play in the establishment of whether or not a particular technology and its effect on a specific patient population furthers the spectrum of evidence-based medicine for patients who have a specific disease or condition. This finding is congruent with the notion that screening technologies, when used appropriately, may expand treatment options for patients; however, at what cost is not clear.

This thematic finding is also supported in the websites that were reviewed. In particular, the AHRQ, BCBSATEC, National Library of Medicine, State of Washington Health Technology Assessment, CTAF, and NICE websites all contain information related to the core mission of the role of technology assessment in improving patient care and impacting cost. AHRQ, BCBCATEC, State of Washington, CTAF, and NICE set

[82] *Ibid,* page 26.

[83] Tatsioni, page 1048.

[84] Rogowski, page 350.

forth criteria that their organizations follow to assess healthcare technology. Similarly, the resources found on the AHRQ and the National Library of Medicine websites support the finding that the purpose of technology assessment is to guide the development of evidence-based medicine in order to increase quality of care while decreasing illness and healthcare costs. Further, AHRQ, BCBSATEC, State of Washington, CTAF, and NICE have specific reviews of technologies that address various screening technologies.

This second thematic finding is also covered in the websites for CMS, AHRQ, BCBS, State of Washington, CTAF and NICE. As discussed above in the Results section, the mission of these organizations is to establish with regard to particular technologies whether the benefit to patient care through the use of the technology might result in inappropriate care and increased cost with marginal return. The technology assessments of specific technologies help the organization, and providers of healthcare for the organization, to understand whether the technology is beneficial or not. Thus, because several organizations have been established to address the concerns of this particular finding, it can be concluded that the current knowledge supports this second finding. However, to what degree is not clear.

Finding 3. Whether patient experience should be a criterion in the HTA process of evaluating a new screening product is not considered.

In both the literature and the website research, there were no statements about particular levels of evidence or criteria that are specific for the evaluation of new technologies that are used to screen for a particular disease. Four articles: Schneider, Birch, Inadomi and Brawley, suggested that patient preferences should be taken into account in evaluating the overall efficacy of a particular intervention. Specifically, Inadomi discusses in great detail that a screening technology might yield impressive

ability to screen for disease; however, if a patient does not adhere to getting the test, then the utility of the technology is substantially diminished.[85] Patient adherence was not examined on the organizational websites that were examined.

As Dr. Goodman suggested in his telephone interview, the selection of a particular technology and a comparison across technology assessment organizations might yield evidence of different criteria. However, Dr. Goodman expressed his concern that the criteria might be different across organization due to exogenous factors such as the mission of the organization (e.g., source of public policy or return on investment for investors), source of funding (e.g., taxpayer funds or private funds), and patient population served (e.g., Medicare or private payer). For example, patient adherence or minimization of cost drivers, might be a mission of one of the organizations and not the others. Thus, while the current knowledge supports this third thematic finding, the rationale for the finding may be difficult to discern due to the influence of outside factors.[86]

Criteria for a screening HTA should be different than therapeutic options because screening technologies serve a different purpose. The purpose of screening tests and procedures is to identify the presence of disease. It is not the purpose of a screening product to actually treat disease. Screening occurs at the beginning of the clinical pathway. To hinge the acceptance of a screening product on the eventual outcome of the case at the end of the clinical pathway may not be appropriate for these technologies. Similarly, a HTA that only evaluates outcomes evidence could be extremely damaging to the development of new screening products because acquiring the level of evidence

[85] Inadomi, page 199.

[86] Teleconference with Dr. Clifford Goodman, April 13, 2009.

necessary for a therapeutic option may not be applicable to a screening product. In addition, if HTA criteria use the highest standards of evidence such as outcomes evidence produced from a randomized controlled clinical trial, innovation in the field of screening may be substantially stifled due to the high standard of evidence required.

Finding 4: Danger of Further Fragmenting Healthcare Policy without Consistent Guidance.

The fourth theme identified in the literature was a risk of further fragmenting healthcare policy in conducting HTA of screening technologies if standards are not consistently applied. While healthcare policy is already fragmented, the development of best practices and standards helps to provide clinical guidance and prevent further fragmentation. Because there are many different agencies conducting HTA, these agencies have different missions, visions, and funding. Vermeulen stated that involving multiple authorities with different agendas results in fragmented policy is a solid conclusion that is a warning sign for the need to develop congruencies between HTA agencies.[87] Further, AHRQ is trying to centralize and unify the development of healthcare policy regarding HTA via their network of EPCs that conduct assessments. However, this point emphasizes that if agencies do not have a point of congruence in the level of evidence required for new technologies, there is a substantial risk of exacerbating the existing fragmentation in healthcare policy by neglecting to take advantage of opportunities to creating consistent standards to guide decision-making.

[87] Vermeulen, V., Coppens, K., & Kesteloot, K. (2001). Impact of health technology assessment on preventive screening in Belgium: Case studies of mammography in breast cancer, PSA screening in prostate cancer, and ultrasound in normal pregnancy. *International Journal of Technology Assessment in Health Care, 17*(3), 316-328.

Quality of Studies

The quality of the studies reviewed in the literature was strong. Although there were only 12 sources to review, each had strong sample sizes, viable study designs, and were specific in the topic each article addressed. With regard to this last point, while none of the articles expressly addressed the research question for this literature review, each paper did make clear what particular topic was to be addressed in the article and engaged in a rigorous review process.

The content in the websites that were reviewed were stronger than the articles. Because websites are created by organizations with specific missions and agendas, biases about the importance or significance of technology assessment could be skewed. However, the strength of the websites can be attributed to their overall purpose: to educate the public about the purpose of technology assessment for their particular organization. As such, the quality of the websites in explaining their purpose and criteria for technology assessment was strong. However, there was lower usefulness for addressing the particular research question for this literature review. Specifically, absent an intensive review of a particular technology assessment for a single product that addresses screening for a particular disease, it is not clear whether these organizations have different review criteria for products that are therapeutic. Further, as described by Dr. Goodman, even an intensive comparison of the review of a technology may not deliver consistent results about the criteria used due to the outside factors that might have influence over the organization.[88]

[88] Teleconference with Dr. Clifford Goodman, April 13, 2009.

Gaps in Current Literature

The clear gap in the current literature is that there is nothing that directly addresses whether the technology assessment criteria for products used to screen for disease are any different than those used to treat a disease. As described above, a review of a specific screening product across technology assessment organizations might provide insight into this question. However, this additional research must include a discussion regarding external factors that might influence the criteria used by the organization such as the mission of the organization, funding source, and patient populations served by the organization.

Implications for Future Research

The objective of this literature review is to determine how healthcare technology agencies assess technologies used to screen for a particular disease. Based on an initial review of the peer-reviewed literature it appears that there is a lack of knowledge and congruence concerning healthcare technology assessment criteria for products used to screen for disease. While there were no articles that were directly addressing the research question, there were several articles that addressed pieces of the issue. These findings may indicate that:

(1) determining how healthcare technology agencies assess technologies used to screen for a particular disease is an area that requires the piecing together of several studies via the information published on the organizational websites identified through the literature review process;

(2) a review of aspects of healthcare, such as drivers of healthcare costs, may result in a more focused view regarding how assessments might be applied to improve screening for disease through cost management and patient quality of care; and

(3) an understanding of outside factors that might influence a technology assessment organization may provide insight into the rationale for the criteria used for a particular technology assessment.

These are areas for recommended additional research in the process of determining whether technology assessment criteria might vary for a screening technology, and what factors might influence the differences.

Limitations of Review Process

The most significant limitations of the review process were whether the keywords used and the databases searched yielded the most relevant and comprehensive sources available to address the research question. Few relevant articles were identified using Google Scholar and PubMed. Similarly, the websites that were researched did not provide specific guidance regarding the research question. Thus, this method of article identification poses a significant risk of missing relevant literature and website resources. Interviews with thought leaders may help in the identification of additional sources of information, policies, and reports.

The second significant limitation of the review process is that the small number of studies and websites may not be indicative of the available information. The keywords used were very specific to "healthcare technology assessment" and "diagnostic" and "screening". If a specific product used to diagnose or screen for a particular disease is used to narrow the search further, it may be that additional resources will become

evident. For example, there are several screening products that have received a technology assessment such as virtual colonoscopy. Interestingly, State of Washington identified this technology as a diagnostic, while CTAF identified it as a screening technology. Regardless of terminology, assessments for the same technology should be compared across different technology assessment agencies to determine if the technology assessment criteria and levels of evidence are congruent or divergent. In addition, international agencies that conduct technology assessment for the same product should be reviewed to determine whether the criteria and levels of evidence are the same or different. From this analysis, perhaps a best practice can be construed to provide guidance to the screening industry regarding the level of evidence that their product must support in order to receive a positive technology assessment recommendation. Thus, perhaps a highest common denominator might be created across countries.

Literature Review Conclusion

Given the rising cost of healthcare and the significant policy discussion regarding the cost of healthcare being driven by varying degrees of patient access to healthcare technologies, it is interesting that there is little data and guidance found regarding the review of products that are used to screen for disease. Arguably, due to the lengthy time to treat many chronic conditions that are discovered by screening products and procedures, one would expect significant data and information about assessing whether or not screening products offer patients better care because the disease is discovered earlier in the clinical pathway. The existing research as discussed here, identifies the preliminary findings that points to gaps in the literature, as well as provides direction regarding initiating additional research for a more thorough understanding of the topic.

It is hoped that the limitations of the current review process can be mitigated through additional research, including interviews with thought leaders. Specifically, the three areas for future research, identified above, will likely reveal additional resources who may further elucidate more of the answer to the research question posed in this study.

Chapter 3: Methods and Analysis Plan

The objective of this study is to determine how health technology assessment agencies assess technologies related to screening products. Specifically, the previous chapter identified four prevalent themes: (1) value of outcomes evidence; (2) evaluation of cost-effectiveness of new technologies; (3) evaluation of patient perspectives; and (4) danger of developing fragmented healthcare policy without guidance.

Study Design and Data Sources

Qualitative interviews involving technology assessment experts, decision-makers, and other personnel at relevant organizations provide the most current and credible information on existing levels of evidence, as well as the future direction of technology assessment for screening products. While quantitative analysis of the levels of evidence required by health technology agencies for diagnostic and screening products would produce a desirable study design, the findings of the literature review indicate that quantitative evidence does not exist for health technology assessment. As a result, qualitative methods are the most appropriate for this study.

This study employs primarily qualitative research methods. Specifically, a "best practices" document review will examine existing levels of evidence employed by technology assessment agencies and other relevant organizations. This approach provides the researcher with an overview of the current methods for health technology assessment for screening products and preliminary evidence of any future trends.

Sources for the document review consist of publicly available technology assessment decisions, working documents from subject matter experts, articles in peer-reviewed journals, newspaper articles, websites, and other relevant documents.

In addition, the study includes interviews involving subject matter experts (e.g., technology assessment experts and decision-makers, etc.) at health insurance companies and other organizations implementing healthcare technology assessments. The basic objectives of these interviews will be to develop a more detailed understanding of the objectives, drivers, structure, effectiveness, organizational commitment to and lessons learned from the development of standards of evidence for screening products. Respondents for the interviews include technology assessment thought leaders, decision makers at technology assessment organizations, or related personnel at organizations implementing technology assessments.

The data collection and analysis portions of the study are estimated to require approximately five months. The limitations presented by this approach lie in the possible absence of standards of evidence for evaluating these types of products, inability for thought leaders to comment on levels of evidence currently in use, and the availability of experts to respond to the interview questionnaire.

Other limitations to this approach include the following: the possible lack of availability of documentation or information on best practices, as they may be considered confidential or proprietary trade secrets, or they may not be publicly accessible; the quality of the information shared (e.g., accuracy or completeness of the documentation); potential bias that interviewees may inject into their responses, either due to presence of the researcher or to the respondents' role in the program being evaluated; and the natural

variability in respondents' abilities to fully perceive and to effectively communicate all of the information requested by the researcher. The researcher's role as a healthcare consultant regarding the coverage and payment of new technologies, which may inhibit the ability of experts to answer a questionnaire or interview, is also a limitation to be considered.

Addressing these limitations requires careful validation techniques such as triangulating multiple sources of information, asking respondents to review and affirm or modify transcripts or other summaries of their interviews.

Data Collection Procedures

Document Review

Technology assessments will be reviewed to select specific products that have been reviewed by HTA agencies. The document review will serve two purposes. First, it will provide a secondary data source for the subset of research questions focused on specific screening technologies. Secondly, it will identify organizations to tap for interviews. Information gathered in the document review will be summarized thematically, in tabular format.

Interviews

Respondents for the interviews will include technology assessment thought leaders, decision makers at technology assessment organizations, or related personnel at organizations implementing technology assessments. Respondents will be identified on the basis of the following factors:

1. A determination, via the best practices and literature review, that the subject's organization is implementing/has implemented (or is a leader in the community that is the intended beneficiary of) a program to evaluate screening technologies.
2. The subject's responsibilities related to either the implementation or ongoing oversight of such program(s).
3. A representative from the manufacturer of the selected screening technology that is being evaluated.
4. A representative from a patient advocacy group that is a supporter or detractor of the screening technology that is being evaluated.

All interviews will be conducted by telephone. See Appendix 1 for the recruitment document and Appendix 2 for the interview guide. Results will be evaluated for completeness and summarized in a data table.

Data Management and Analysis

Since the research question is an emerging field of technology assessment, it is anticipated that relatively few interviews, 8 to 10, will be conducted. Considerations regarding the bias of the interviewed subject and agency must be taken into account. The analysis will likely parallel that of a literature review, e.g., groupings of themes, a discussion of findings, and a presentation of conclusions. The notes taken from the interviews will be analyzed to identify themes and to compare and contrast responses across interviews. To the extent possible, the data will be quantified. For example, if similar technology review endpoints between agencies are identified, these themes will be counted or weighted either by frequency of mention, extent of treatment of a theme, or

both. Further, themes illuminated by the initial document review may yield a set of scales with which to quantify one or more aspects of the interviews.

If an organization has performed a quantitative analysis of their program, and is willing to share that information, it will also be included in the study as secondary data. For example, there may be evidence tables or criteria that have already been developed by an organization to evaluate screening products. If this information is made available to the researcher, it will be included in the study. However, issues related to bias and validity must be considered if these materials are not published or peer-reviewed.

Proposed Study Deliverables

This research will result in a detailed summation of current methods employed by technology assessment agencies designed to accomplish any of the following related to identifying the criteria used to evaluate screening products: (1) value and of outcomes evidence; (2) evaluation of cost-effectiveness of new technologies; (3) evaluation of patient experience; and (4) danger of further fragmenting healthcare policy when there may be an opportunity to create a standard.

The approach outlined above includes an intentional narrowing of the research focus from very broad to the most narrow. This method both anticipates that the levels of evidence identified and examined will drastically decline as the focus narrows, and appreciates that there will be valuable information to be derived from each program that shares information about how they conduct technology assessments for screening products.

The summation will include a discussion and evaluation of any and all relevant data or other evidence provided by the evaluated programs that indicates success against

stated objective(s). Examples of such markers include: (1) stated levels of evidence for screening products; (2) the thoroughness of the health technology assessments for a particular technology; and (3) any information regarding the future development of specific guidelines in this area of health technology assessment.

A recommendation will be created for health technology assessment agencies, health insurers, and other organizations to take to establish levels of evidence for the technology assessment of screening products in the United States. Depending upon the volume and quality of information discovered through the research, this may consist of simply of a "best practices" checklist in the event that few programs or methods exist to be reviewed, or as elaborate as a specific description of relevant levels of evidence to integrate into a health technology assessment to guide an organization through the prioritization and sequencing of levels of evidence to effectively evaluate new screening products. In addition, where relevant, differing objectives of technology assessment panels will be acknowledged.

IRB and Confidentiality Issues

This study will require research that involves direct interaction with human subjects; therefore all relevant information was submitted to the Public Health-Nursing Institutional Review Board (IRB). The study was approved by the IRB and is identified as Study Number 11-0624.

Qualifying subjects will be recruited via email and telephone contact. Respondents will be provided with a full written description of the study prior to the interviews, and provided with an opportunity to ask questions and/or express concerns via the method of their choice (e.g., email or telephone) prior to scheduling their initial

interview. Verbal or signed consent from proposed respondents will be obtained prior to any data collection, in accordance with the IRB approval conditions.

Appointments will be made ahead of time for the phone interviews. This will give the interviewees opportunities to schedule a time when they can talk in a private space. Voice recording will not be used. At no point will the name, location or any other element that may allow the reader to ascertain the specific insurer or organization be identified. Maximum interview time will be 60 minutes.

Chapter 4: Health Technology Assessment Comparison
Computed Tomographic (CT) Colonography for Colon Cancer Screening

Computed Tomographic (CT) colonography, also called virtual colonoscopy, is a screening technology for colon cancer. A review of the HTA websites set forth in Table 10 identified CT colonography as the only screening technology with publicly available HTAs from different HTA agencies.

Table 10. HTA Websites

Institution/Agency	Website	HTA of CT Colonography
Centers for Medicare and Medicaid Services	http://www.cms.hhs.gov/default.asp?	No
Agency for Healthcare Research and Quality	http://www.ahrq.gov/	No
Blue Cross Blue Shield Association Technology Evaluation Center	http://www.bcbs.com/blueresources/tec/	Yes
California Technology Assessment Forum	http://www.ctaf.org/	Yes
U.S. Preventative Services Task Force	http://www.ahrq.gov/CLINIC/uspstfix.htm	Yes
National Library of Medicine – National Institutes of Health	http://www.nlm.nih.gov/	No
State of Washington's Health Technology Assessment	http://www.hta.hca.wa.gov/	Yes
Evaluation of Genomic Applications in Practice and Prevention	http://www.egappreviews.org/	No
Cochrane Collaboration	http://www.cochrane.org/	No
National Institute of Health and Clinical Excellence	http://www.nice.org.uk/	Yes

Background: Colorectal Cancer

Colon cancer forms in the longest part of the large intestine in the tissues of the colon. Rectal cancer forms in the last several inches of the large intestine in the tissues of the rectum. Colorectal cancer is a disease in which cells in the colon or rectum become

abnormal and divide without control, forming a mass that may spread to form new masses in other parts of the body.[89] According to the National Cancer Institute:

> "Colorectal cancer is the third most common type of non-skin cancer in men (after prostate cancer and lung cancer) and in women (after breast cancer and lung cancer). It is the second leading cause of cancer death in the United States after lung cancer. Although the rate of new colorectal cancer cases and deaths is decreasing in this country, more than 145,000 new cases were diagnosed and more than 49,000 people died from this disease each year over the past 5 years."[90]

In 2010, there are an estimated 102,900 new cases of colon cancer and 39,670 of rectal cancer in the United States.[91] Deaths from colon cancer and rectal cancer, combined, in the United States are 51,370 in 2010.[92]

The United States Preventative Services Task Force (USPSTF) reports that more than 80% of diagnosed cases of colorectal cancer occur in patients over age 55. USPSTF also reports the age-adjusted incidence is 51.6 per 100,000 persons, with a lifetime risk of diagnosis of 5.7% for men and 5.1% for women.[93] Increased incidence is associated with increased age, male sex and black race. The incidence rate has decreased over the past 20

---

[89] National Cancer Institute, U.S. National Institutes of Health. Colorectal Cancer Screening. http://www.cancer.gov/cancertopics/factsheet/Detection/colorectal-screening Accessed: November 13, 2010.

[90] *Ibid.*

[91] National Cancer Institute, U.S. National Institutes of Health. Colon and Rectal Cancer. http://www.cancer.gov/cancertopics/types/colon-and-rectal Accessed: November 13, 2010.

[92] *Ibid*.

[93] U.S. Preventative Services Task Force. (2008). Screening for colorectal cancer: U.S. Preventative Services Task Force Recommendation. *Ann Intern Med*, 149:627-37, at 631.

years among men of all racial and ethnic groups except for Hispanics/Latinos and Alaska Natives. Incidence has stabilized among women of all racial and ethnic groups except Alaska Natives.[94] In 2002, USPSTF issued a strong recommendation that average-risk adults, age 50 or older, receive screening for colorectal cancer.[95]

Colorectal Cancer Screening

Although the exact causes of colorectal cancer are not known, there are studies that have identified factors, such as age, presence of colon or rectal polyps, and personal history; that are linked to an increased likelihood of developing the disease.[96] The early identification and removal of polyps may help to prevent colorectal cancer. As a result, screening for colorectal cancer can help to detect cancer, polyps, and nonpolypoid lesions that might develop into colorectal cancer over time. Colorectal cancer is more treatable when found early. If the screening reveals an issue, diagnosis and treatment can occur quickly.[97]

There are data from several sources that demonstrate a reduction in mortality when there is regular screening for colorectal cancer.[98] However, "[d]espite its

---

[94] *Ibid.*

[95] *Ibid.*

[96] National Cancer Institute, U.S. National Institutes of Health. Colorectal Cancer Screening. http://www.cancer.gov/cancertopics/factsheet/Detection/colorectal-screening Accessed: November 13, 2010.

[97] *Ibid.*

[98] Johnson CD, Chen MH, Toledano AY, et al. (2008). Accuracy of CT colonography for detection of large adenomas and cancers. *N Engl J Med*, 359(12):1207-17, page 1208. Cross-referencing the following sources:

Liedenbaum MH, Vries AH, Rijn AF, et al. (2010). CT colonography with limited bowel preparation for the detection of colorectal neoplasia in an FOBT positive screening population. *Abdominal Imaging*, 35:6, 661-668.

effectiveness, colorectal-cancer screening remains underused for many reasons, including drawbacks in terms of the performance, comfort, availability, and expense of currently endorsed test options."[99] Table 11 sets forth the National Cancer Institute's catalog of advantages and disadvantages of the currently available methods for colorectal cancer screening.

Table 11. National Cancer Institute: Advantages and Disadvantages of Colorectal Cancer Screening Tests[100]

Test	Advantages	Disadvantages
Fecal Occult Blood Test (FOBT)	• No cleansing of the colon is necessary. • Samples can be collected at home. • The cost is low compared with other colorectal cancer screening tests. • FOBT does not cause bleeding or tearing/perforation of the lining of the colon.	• This test fails to detect most polyps and some cancers.[101,102] • False-positive results (the test suggests an abnormality when none is present) are possible.[103] • Dietary restrictions and changes, such as avoiding meat, certain vegetables, vitamin C, iron supplements, and aspirin, and increasing fiber consumption, are often recommended for several days before a guaiac FOBT. These restrictions and changes are not required for immunochemical FOBT. • Additional procedures, such as colonoscopy, may be necessary if the test indicates an abnormality.

Heresbach D, Chauvin P, Grolier J, et al. (2010). Cost-effectiveness of colorectal cancer screening with computed tomography colonography or fecal blood tests. *European Journal of Gastroenterology & Hepatology*, 22:11, 1372-1379.

Sacher-Huvelin S, Coron E, Gaudric M, et al. (2010). Colon capsule endoscopy vs. colonoscopy in patients at average or increased risk of colorectal cancer. *Alimentary Pharmacology & Therapeutics*, 32:9, 1145-1153.

[99] *Ibid.*

[100] National Cancer Institute, U.S. National Institutes of Health. Colorectal Cancer Screening. http://www.cancer.gov/cancertopics/factsheet/Detection/colorectal-screening Accessed: November 13, 2010.

[101] Burch JA, Soares-Weiser K, St John DJ, et al. (2007). Diagnostic accuracy of faecal occult blood tests used in screening for colorectal cancer: A systematic review. *Journal of Medical Screening*; 14(3):132–137.

[102] Ouyang DL, Chen JJ, Getzenberg RH, Schoen RE. (2005). Noninvasive testing for colorectal cancer: A review. *American Journal of Gastroenterology*; 100(6):1393–1403.

[103] *Ibid*, footnotes 14 and 15.

Test	Advantages	Disadvantages
Sigmoidoscopy	• The test is usually quick, with few complications. • For most patients, discomfort is minimal. • In some cases, the doctor may be able to perform a biopsy (the removal of tissue for examination under a microscope by a pathologist) and remove polyps during the test, if necessary. • Less extensive cleansing of the colon is necessary with this test than for a colonoscopy.	• This test allows the doctor to view only the rectum and the lower part of the colon. Any polyps in the upper part of the colon will be missed. • There is a very small risk of bleeding or tearing/ perforation of the lining of the colon.[104] • Additional procedures, such as colonoscopy, may be necessary if the test indicates an abnormality.
Colonoscopy	• This test allows the doctor to view the rectum and the entire colon. • The doctor can perform a biopsy and remove polyps or other abnormal tissue during the test, if necessary.	• This test may not detect all small polyps, nonpolypoid lesions, and cancers, but it is one of the most sensitive tests currently available. • Thorough cleansing of the colon is necessary before this test. • Some form of sedation is used in most cases. • Although uncommon, complications such as bleeding and/or tearing/perforation of the lining of the colon can occur.[105]
Virtual Colonoscopy	• This test allows the doctor to view the rectum and the entire colon. • This is not an invasive procedure, so there is no risk of bleeding or tearing/perforation of the lining of the colon.	• This test may not detect all small polyps, nonpolypoid lesions, and cancers.[106, 107] • Thorough cleansing of the colon is necessary before the test. • If a polyp or nonpolypoid lesion 6 to 9 millimeters in size or larger is detected, standard colonoscopy, usually immediately after the virtual procedure, will be recommended to remove the polyp or lesion or perform a biopsy.[108, 109]

[104] Gatto NM, Frucht H, Sundararajan V, et al. (2003). Risk of perforation after colonoscopy and sigmoidoscopy: A population-based study. *Journal of the National Cancer Institute*; 95(3):230–236.

[105] *Ibid.*

[106] Pickhardt PJ, Choi JR, Hwang I, et al. (2003). Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *New England Journal of Medicine*; 349(23):2191–2200.

[107] Johnson CD, Chen MH, Toledano AY, et al. (2008). Accuracy of CT colonography for detection of large adenomas and cancers. *New England Journal of Medicine*; 359(12):1207–1217.

[108] Rex DK, ACG Board of Trustees. (2004). American College of Gastroenterology action plan for colorectal cancer prevention. *American Journal of Gastroenterology*; 99(4): 574–577.

[109] Summerton S, Little E, Cappell MS. (2008) CT colonography: Current status and future promise. *Gastroenterology Clinics of North America*; 37(1):161–189.

Test	Advantages	Disadvantages
Double Contrast Barium Enema (DCBE)	• This test usually allows the doctor to view the rectum and the entire colon. • Complications are rare. • No sedation is necessary.	• This test may not detect some small polyps and cancers.[110] • Thorough cleansing of the colon is necessary before the test. • False-positive results are possible. • The doctor cannot perform a biopsy or remove polyps during the test. • Additional procedures are necessary if the test indicates an abnormality.
Digital Rectal Exam (DRE)	• Often part of a routine physical examination. • No cleansing of the colon is necessary. • The test is usually quick and painless.	• The test can detect abnormalities only in the lower part of the rectum. • Additional procedures are necessary if the test indicates an abnormality.

The American Cancer Society recommends that people at average risk should begin screening for colorectal cancer at age 50. Options for testing are as follows:[111]

Tests that find polyps and cancer:

- Flexible sigmoidoscopy every 5 years
- Colonoscopy every 10 years
- Double-contrast barium enema every 5 years
- CT colonography (virtual colonoscopy) every 5 years

Tests that mainly find cancer:

- Fecal occult blood test (FOBT) every year
- Fecal immunochemical test (FIT) every year
- Stool DNA test (sDNA), interval uncertain

An interval of 10 years is recommended for rescreening patients with a negative finding from a colonoscopy. This interval "is based on the average time to progression

[110] PDQ® Cancer Information Summary. National Cancer Institute; Bethesda, Maryland. *Colorectal Cancer Screening—Health Professional*. Date last modified: 08/26/2008. Available at http://www.cancer.gov/cancertopics/pdq/screening/colorectal/healthprofessional. Accessed: October 3, 2008.

[111] American Cancer Society. Can colorectal cancer be found early? http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-detection Accessed: December 27, 2010.

from an adenomatous polyp to cancer is at least 10 years."[112] The Centers for Disease Control and Prevention states that "[t]he decision to be screened after age 75 should be made on an individual basis."[113]

USPSTF notes that there are harms associated with all forms of colorectal cancer screening. For example, a positive test leads to follow up testing to confirm the diagnosis, usually a colonoscopy. Colonoscopy can result in: morbidity, patient anxiety, inconvenience, discomfort, as well as additional medical costs.[114] Thus, while screening is recommended for age-appropriate asymptomatic patients to decrease mortality, there is no preferred method and there is risk of harm with each option.

CT Colonography for Colorectal Cancer Screening

The impetus for CT colonography is to allow visualization of the colon and rectum with lower risk of perforation and other complications.[115] CT colonography uses helical computed tomography combined with computer software to generate detailed images of the inner surface of the colon for review by a radiologist.[116] Similar to a conventional colonoscopy, CT colonography requires a full bowel cleansing and uses a tube to fill the colon with air. The test typically takes 10 minutes and involves two scans.

---

[112] Madison Clinic, Harborview Medical Clinic, University of Washington Medical Center. Screening for Colorectal Cancer. http://depts.washington.edu/madclin/providers/guidelines/colorectal.html Accessed: December 27, 2010.

[113] Centers for Disease Control and Prevention. Colorectal Cancer Screening Guidelines. http://www.cdc.gov/cancer/colorectal/basic_info/screening/guidelines.htm Accessed: December 27, 2010.

[114] U.S. Preventative Services Task Force. (2008). Screening for colorectal cancer: U.S. Preventative Services Task Force Recommendation. *Ann Intern Med*, 149:627-37, at 632.

[115] *Ibid,* at 633.

[116] CT Colonography ("Virtual Colonoscopy") for Colon Cancer Screening. (2009). Blue Cross Blue Shield Association, Technology Evaluation Center. Assessment Program, Volume 24, No. 1, page 1. August 2009.

One while the patient is lying on their back and the other with the patient lying on their stomach. Each scan lasts 10 to 15 seconds.[117] Advantages of the technology include: fast imaging of the entire colorectum, minimally invasive technique, no need for sedation, and there is a low-risk for complications from the procedure.[118]

Prior to 2008, several small clinical trials for CT colonography had been conducted. In 2008, the results of the National CT Colonography Trial of the American College of Radiology Imaging Network (ACRIN study) were published in the *New England Journal of Medicine*.[119] The objective of the ACRIN study was to evaluate the technology in a prospective trial setting in order to evaluate the performance of CT colonography versus the standard of care colonoscopy. The ACRIN study involved 15 training centers and recruited 2600 asymptomatic patients aged 50 years old or older. "CT colonographic images were acquired with the use of standard bowel preparation, stool and fluid tagging, mechanical insufflation, and multidetector-row CT scanners (with 16 or more rows). Radiologists trained [and tested for participation this trial] in CT colonography reported all lesions measuring 5mm or more in diameter…[t]he primary end point was detection by CT colonography of histologically confirmed large adenomas and adenocarcinomas (10 mm in diameter or larger) that had been detected by

---

[117] American Cancer Society. Can colorectal cancer be found early? http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-detection Accessed: December 28, 2010.

[118] Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. *New England Journal of Medicine* 2008; 359(12):1207–1217.

[119] *Ibid.*

colonoscopy; detection of smaller colorectal lesions (6 to 9 mm in diameter) were also evaluated."[120]

The ACRIN study is of particular interest because completed data were collected for 2531 patients (97%) and same day CT colonographic and colonoscopic exams were performed on 2512 of the 2531 patients.[121] In addition, 89% of participants had no known risk factors for colorectal cancer, other than age. This multicenter study of asymptomatic patients showed that 90% of patients with adenomas and cancers measuring 10 mm or more were identified with CT colonography. As a result, the ACRIN study substantially supported previous studies regarding the accuracy of CT colonography.[122]

Health Technology Assessment (HTA) and CT Colonography

Screening technologies are difficult for HTA agencies to review. In the assessment process, HTA agencies are looking for screening technologies that can both: (1) identify disease with sufficient sensitivity and specificity to be considered accurate; and (2) allow available screening technology interventions to be used in order to avoid the shortening of life and reduction of functioning status before other diseases and aging do so. As a result, the proof of the value of a screening technology depends to a much greater degree than other medical technologies on: (1) the availability of successful interventions to patients; and (2) the ability to counter hypotheses that a large proportion of positive findings are either indolent or sufficiently slow in progression to mean that other conditions are more likely to cause mortality of morbidity.

---

[120] *Ibid.*

[121] *Ibid.*

[122] *Ibid.*

These challenges related to the assessment of screening technologies are apparent in the reports issued by five different HTA agencies regarding CT colonography. Assessments and reports were issued from 2004 to 2009 and resulted in two agencies finding sufficient evidence to support the use of this technology. The agencies, dates of review and outcome of the HTA are set forth in Table 12.

Table 12. HTA Agencies with Assessments of CT Colonography

Date	HTA	Outcome
June 2004	California Technology Assessment Forum	Insufficient Evidence
July 2004	Blue Cross Blue Shield Association TEC	Insufficient Evidence
June 2005	National Institute of Health and Clinical Excellence	Sufficient Evidence
August 2008	State of Washington HTA	Insufficient Evidence
November 2008	U.S. Preventative Services Task Force	Insufficient Evidence
March 2009	California Technology Assessment Forum	Insufficient Evidence
August 2009	Blue Cross Blue Shield Association TEC	Sufficient Evidence

Detail follows regarding each technology assessment agency, levels of evidence applied to CT colonography, and the outcome for each assessment.

California Technology Assessment Forum

California Technology Assessment Forum employs a four-step assessment process for the evaluation of new and emerging technologies.

Step 1.
Literature Search
↓
Step 2.
Data Analysis & Recommendation
↓
Step 3.
Public Meetings & Invitations
↓
Step 4.
Panel Member Vote

First, a scientific literature search is conducted and may be augmented by soliciting positions and opinions from medical specialty societies and physician experts. In addition, reports from government agencies are obtained. Examples of government agencies are the U.S. Food and Drug Administration, National Institutes of Health, and Agency for Healthcare Research and Quality.[123]

Second, data from the first step are analyzed and a recommendation is formed. Medical consultants who are experts in evaluating scientific evidence prepare a recommendation. The Blue Shield of California Foundation contracts with the University of California, San Francisco (UCSF) to conduct the evidence review, analysis, and prepare a written recommendation. In order to be considered safe and effective, the technology must meet five technology assessment criteria.[124] The five criteria

[123] Assessment Process. California Technology Assessment Forum. http://www.ctaf.org/section/assessment/process Accessed: November 21, 2010.

[124] *Ibid.*

are set forth in Table 13.[125]

Table 13. CTAF HTA Criteria

No.	Criteria
1	The technology must have final approval from the appropriate government regulatory bodies.
2	The scientific evidence must permit conclusions concerning the effect of the technology regarding health outcomes.
3	The technology must improve net health outcomes.
4	The technology must be as beneficial as any established alternatives.
5	The improvement must be attainable outside the investigational setting.

Regarding the second criterion, evidence of effectiveness is graded as follows:[126]

- **Level 1:** Randomized trials that had enough power to demonstrate a statistically significant health outcome.
- **Level 2:** Randomized trials with results that were not statistically significant but where a larger trial might not have shown a clinically important difference.
- **Level 3:** Nonrandomized concurrent cohort comparisons between contemporaneous patients.
- **Level 4:** Nonrandomized historical cohort comparisons between current patients and former patients (from the same institution or from the literature).
- **Level 5:** Case series without control subjects.

While Level 1 is the preferred to determine whether the second criterion is met, in the absence of this type of study, "technologies may meet this criterion if, overall, Level 2-4 studies indicate that:

a. The technology provides substantial benefits to important health outcomes, and
b. The new technology has been shown to be safer or more beneficial than existing technologies or alternative treatments in comparative studies.

In general, technologies will not be approved based on evidence from Level 5 studies (case studies without controls)."[127]

---

[125] Criteria. California Technology Assessment Forum. http://www.ctaf.org/section/assessment/criteria Accessed: November 21, 2010.

[126] *Ibid.* Evidence about the effectiveness are according to the criteria proposed by: Cook DJ, et al. (1992). Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*. 4(suppl):305S-311S.

Third, CTAF holds open forum meetings three times a year. "Invitations are extended to leaders of local medical and professional societies, representatives of healthcare and regulatory agencies, physician experts from academic centers and the community, health plan representatives and manufacturers. Guests and medical professionals are invited to attend and comment at the meeting."[128]

Fourth, once a topic is presented, reviewed, and public opinion has been provided, then the Forum panel takes a vote. The panel may vote to accept, modify, reject, or table the discussion until more information is available.[129]

The first HTA for CT colonography occurred in June 2004.[130] The second occurred in March 2009, after the publication of the ACRIN study data.[131] Both times, CTAF recognized that the technology met the evidence criteria for 1 and 2; however, ultimately rejected the technology due to an inability to satisfy criteria 3, 4, and 5:

> ***Criterion 1: The technology must have final approval from the appropriate government regulatory bodies.***

CTAF noted that the technology does have regulatory clearance from the U.S. Food and Drug Administration. As a result, in both the 2004 and 2009 HTAs, criterion 1 is met.

---

[127] Criteria. California Technology Assessment Forum. http://www.ctaf.org/section/assessment/criteria Accessed: November 21, 2010.

[128] Assessment Process. California Technology Assessment Forum. http://www.ctaf.org/section/assessment/process Accessed: November 21, 2010.

[129] *Ibid.*

[130] Feldman M. (2004). Computed tomographic colonography (virtual colonoscopy) for screening of colorectal cancer. *California Technology Assessment Forum*. June 9, 2004.

[131] Walsh J. Computed tomographic colonography (virtual colonoscopy) for colorectal cancer screening in average risk individuals. (2009). *California Technology Assessment Forum*. March 11, 2009.

Criterion 2: The scientific evidence must permit conclusions concerning the effectiveness of the technology regarding health outcomes.

CTAF's 2004 HTA indicates that colorectal screening is effective based on results from a randomized controlled study of FOBT and case-control studies of sigmoidoscopy. As a result a decision was made that CT colonography did demonstrate that it has equal or better performance to the "gold standard" of optical colonoscopy.[132] The data supported CTAF's level of evidence 3, nonrandomized concurrent cohort comparisons between contemporaneous patients; and CTAF's level 5, case series without control subjects. As a result, criterion 2 was met.[133]

In CTAF's 2009 HTA, the ACRIN study data is considered as a part of satisfying this criterion. CTAF concluded that "[t]he results of this study suggest that the accuracy of CT [colonography] in detecting significant colorectal abnormalities is relatively comparable to [colonoscopy]."[134] The data represents CTAF level 3, nonrandomized concurrent cohort comparisons between contemporaneous patients. Criterion 2 was met.[135]

Criterion 3: The technology must improve the net health outcomes.

In CTAF's 2004 HTA, this third criterion reads: "[t]he technology must improve the net health outcome. For diagnostic tests, there is evidence that the use of the test

---

[132] Feldman M. Computed tomographic colonography (virtual colonoscopy) for screening of colorectal cancer. (2004). *California Technology Assessment Forum*. June 9, 2004. Page 5.

[133] *Ibid.*

[134] Walsh J. Computed tomographic colonography (virtual colonoscopy) for colorectal cancer screening in average risk individuals. (2009). *California Technology Assessment Forum*. March 11, 2009. Page 8.

[135] *Ibid, page 9.*

would result in improved medical management in a way that will benefit the patient."[136] The HTA noted that of the over 20 studies comparing the accuracy of CT colonography to colonoscopy, only 5 contain data with patients from an average risk population who present for routine screening.[137] As a result, the efficacy of the test in an average-risk screening population was called into question. In addition, there was insufficient published evidence to establish that CT colonography could be reliably taught and implemented in a community-based practice.[138] Further, the HTA noted that while CT colonography might have better patient acceptance, data has only been examined in a few studies and if pre-procedure preparation for colonoscopy was less rigorous, it could be that colonoscopy could become better tolerated. As a result, improved medical management is not clear and the criterion was deemed not met.[139]

CTAF's 2009 HTA specifically examined average risk individuals.[140] The data from the ACRIN study helped to overcome this barrier from CTAF's 2004 HTA in which only high risk patients were considered. In addition, this third criterion was truncated to: "[t]he technology must improve net health outcomes."[141] The HTA noted three benefits of CT colonography: (1) ability to detect small polyps; (2) fewer complications than colonoscopy; and (3) procedure may be more acceptable to patients than more invasive

---

[136] Feldman M. Computed tomographic colonography (virtual colonoscopy) for screening of colorectal cancer. (2004). *California Technology Assessment Forum*. June 9, 2004. Page 6.

[137] *Ibid.*

[138] *Ibid, page 10.*

[139] *Ibid, page 11 and 12.*

[140] Walsh J. Computed tomographic colonography (virtual colonoscopy) for colorectal cancer screening in average risk individuals. (2009). *California Technology Assessment Forum*. March 11, 2009. Page 6.

[141] *Ibid,* page 9.

techniques.[142] Potential harms included: (1) concern regarding radiation exposure over a patient's lifetime; (2) although relatively few, procedure related harms such as one out of 2531 patients hospitalized for bacteremia for a patient who had both a CT colonography followed by a colonoscopy; and (3) identification of extra-colonic findings. This last concern is related to the CT obtaining images of structures outside the colon. In some cases, radiologists requested follow up tests to investigate the finding and this resulted in increased patient anxiety and overall cost. CTAF found that "[a]lthough it is possible that early detection of these abnormalities may lead to improved outcomes, it is also possible that there will be additional unnecessary medical testing of these abnormalities with associated anxiety."[143] Thus, although the accuracy of CT colonography approaches that of colonoscopy, CTAF concluded that it is unclear whether the harms outweigh the benefits. As a result, this criterion was not met.[144]

Criterion 4: The technology must be as beneficial as any established alternatives.

Both the 2004 and 2009 HTAs noted three established alternatives used for screening for colorectal cancer: (1) Fecal Occult Blood Testing (FOBT); (2) flexible sigmoidoscopy; and (3) colonoscopy. The 2004 HTA noted that because "[m]ost of the current studies of CT colonography do not demonstrate sensitivity comparable to colonoscopy, and none demonstrate a reduction in mortality seen with flexible

[142] *Ibid.*

[143] *Ibid,* page 10.

[144] *Ibid,* page 11.

sigmoidoscopy or FOBT."[145] Further, because there was limited evidence of efficacy in an average risk population, the fourth criterion was not met.[146]

The 2009 HTA applauded the additional evidence published in the ACRIN study. However, the HTA also set forth that the benefit of the technology depends not only on its diagnostic accuracy, but also on the overall risks and benefits as compared to other screening technologies. Concerns related to radiation exposure and the identification of extra-colonic findings led to this criterion not being met.[147]

Criterion 5: The improvement must be attainable outside the investigational settings.

CTAF's 2004 HTA indicated that there was not enough evidence to demonstrate that CT colonography could be disseminated without a great deal of extra training for radiologists. As a result, the clinical data could not be easily replicated in the community based setting and the criterion was not met.[148]

CTAF's 2009 HTA acknowledged the features of the ACRIN study to make the procedure more easily shared outside the investigational setting. However, again, the potential risks related to radiation exposure and high rates of extra-colonic findings were

[145] Feldman M. Computed tomographic colonography (virtual colonoscopy) for screening of colorectal cancer. (2004). *California Technology Assessment Forum*. June 9, 2004. Page 13.

[146] *Ibid,* page 14

[147] Walsh J. Computed tomographic colonography (virtual colonoscopy) for colorectal cancer screening in average risk individuals. (2009). *California Technology Assessment Forum*. March 11, 2009. Pages 11 and 12.

[148] Feldman M. Computed tomographic colonography (virtual colonoscopy) for screening of colorectal cancer. (2004). *California Technology Assessment Forum*. June 9, 2004. Page 14.

cited as not establishing an improvement in health outcomes. As a result, this criterion was not met.[149]

While the ACRIN study changed the scope of CTAF's HTA from 2004 to 2009 to include average risk patients and helped to establish CT colonography as a technology that could be disseminated into community-based practice, concerns regarding the possible harm of the procedure outweighed the benefit. In both HTAs, the technology failed to meet all five criteria and did not receive a positive recommendation.

Blue Cross Blue Shield Association Technology Evaluation Center

Like CTAF, the Blue Cross Blue Shield Association Technology Evaluation Center (TEC) both conducted HTAs of CT colonography on two occasions. Both agencies have a similar five criteria methodology for conducting HTAs of new technologies. TEC's criteria are almost identical to those of CTAF, except the third criterion for CTAF indicates that the technology must improve the net health outcome. TEC's third criterion indicates that the technology must have a net health outcome.

While CT colonography did not receive a positive TEC recommendation when it was first assessed in 2004, it did receive a positive recommendation in 2009 after the issuance of the ACRIN study data. Notably, the authors commented that the 2004 HTA concluded that the TEC criteria were not met for the following reasons: (1) the sensitivity in the literature was variable among studies; (2) variability in performance could be attributed to interpreter experience and technical factors; and (3) that clear criteria were

[149] Walsh J. Computed tomographic colonography (virtual colonoscopy) for colorectal cancer screening in average risk individuals. (2009). *California Technology Assessment Forum*. March 11, 2009. Page 12.

needed regarding polyp size threshold for removal and for frequency of screening in order to appropriately assess the effectiveness of the technology.[150]

The ACRIN study helped to fill the gaps of the prior HTA. The TEC HTA was initially reviewed by the Blue Cross and Blue Shield Medical Advisory Panel on June 10, 2008. Their review was purposefully delayed until the ACRIN study data was published later that year.[151] The primary aim of the study was to evaluate the sensitivity of CT colonography vs. colonoscopy in detecting lesions greater than 10 mm with a secondary aim of detecting polyps from 5-10 mm. Second, the ACRIN trial trained radiologists and tested their proficiency in order to qualify to participate in the study.[152] Thus, the ACRIN study data helped to establish TEC confidence in both the sensitivity and the ability to introduce the technology to the community-based setting.

Criterion 1. The technology must have final approval from the appropriate governmental regulatory bodies.

TEC determined that the appropriate regulatory body, U.S. Food and Drug Administration, had approved CT colonography technology. This criterion was satisfied.[153]

Criterion 2. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.

---

[150] CT colonography ("virtual colonography") for colon cancer screening. (2009). Technology Evaluation Center. Assessment Program. Vol. 24, No. 1, August 2009. Blue Cross Blue Shield Association.

[151] *Ibid, page 10.*

[152] *Ibid, page 2.*

[153] *Ibid, page 21.*

TEC's HTA determined that there is no direct evidence linking CT colonography to health outcomes. However, an "inference of effectiveness" can be made based on the sensitivity data presented in the ACRIN study which demonstrated 90% sensitivity for detection of polyps 10 mm or larger. As a result, this criterion was met.[154] Thus, like CTAF's HTA, while there was no definitive direct evidence of the technology affecting health outcomes, the ACRIN study helped to establish sufficient evidence.

Criterion 3. The technology must improve the net health outcome; and Criterion 4. The technology must be as beneficial as any established alternatives.

TEC's HTA bundled these two criteria together. Again, citing the chain of evidence and the value of the ACRIN study data regarding sensitivity and 86% specificity, the HTA noted that some unnecessary colonoscopies might have to be conducted due to false-positives. However, the benefit of identifying and removing cancer outweighed the unquantifiable risks, presumed to be small, of detecting unrelated health problems and radiation exposure. As such, these criteria were deemed satisfied in the HTA.[155]

Criterion 5. The improvement must be attainable outside the investigational settings.

Again, citing the ACRIN study data, the techniques employed to ensure that radiologists were appropriately trained and had the necessary skill to perform the CT colonography, this criterion was deemed satisfied in the HTA. Specifically, the HTA

---

[154] *Ibid, page 22.*

[155] *Ibid.*

noted that if the outcomes of the ACRIN study could be replicated in the community, health outcomes would indeed be improved.[156]

The ACRIN study data helped satisfy the TEC HTA. Notably, the size of the study, the sensitivity and specificity data, and the ability to train practitioners provided the reviewers with confidence in CT colonography despite the risks identified by both CTAF and TEC.

National Institute of Health and Clinical Evidence

The United Kingdom's National Institute of Health and Clinical Evidence (NICE) issued a guidance regarding CT colonography in June 2005. The two-page guidance indicates that the current safety and efficacy data for CT colonography "appears adequate to support the use of this procedure provided that the normal arrangements are in place for consent, audit and clinical governance."[157] The guidance provides its support to CT colonography both as a diagnostic in symptomatic patients, as well as in high risk asymptomatic patients.

Efficacy was established via sensitivity and specificity data from 14 studies with a total of 1324 patients. Pooled per-patient sensitivity is set forth below in Table 14.

Table 14. Pooled Per-Patient Sensitivity[158]

Polyp Size	Sensitivity	Confidence Interval
10 mm or larger	88%	95% CI, 84-93%
6-9 mm	84%	95% CI, 80-89%
5 mm or smaller	65%	95% CI, 57-73%

[156] *Ibid.*

[157] Computed tomographic colonography (virtual colonoscopy). National Institute for Health and Clinical Excellence. June 2005. http://www.nice.org.uk/guidance/IPG129 Accessed: November 21, 2010.

[158] *Ibid.*

In addition, a large study of 1233 asymptomatic patients reported "per-polyp" sensitivity set forth in Table 15.

Table 15. Pooled Per-Polyp Sensitivity[159]

Polyp Size	Sensitivity	Confidence Interval
10 mm or larger	81%	95% CI, 76-85%
6-9 mm	62%	95% CI, 58-67%
5 mm or smaller	43%	95% CI, 39-47%

While advisors to the guidance noted that small, flat, lesions could be missed; they also stated that these could be missed by other diagnostic techniques as well.[160] While the technology is dependent on the experience of the technician and the type of equipment, the guidance states that the technology could be particularly helpful in detecting tumors in patients who are frail and/or elderly.[161] As a result, NICE supports the use of the technology.

State of Washington HTA

The Washington State Health Technology Clinical Committee (HTCC) consists of 11 health care practitioners. The function of the independent committee is to determine how technologies are covered by state agencies and are based on the technology's safety, efficacy, and cost effectiveness.[162] "Evidence includes a report concerning the technology provided by a company specializing in objective reviews of pertinent scientific literature; information submitted by the affected state agencies; and

[159] *Ibid.*

[160] *Ibid.*

[161] *Ibid.*

[162] Washington State Health Care Authority. (2008). Health Technology Clinical Committee. Findings and Coverage Decision: Computed Tomographic Colonography. Meeting date: February 15, 2008. Final adoption: August 15, 2008. Page 4.

public comment."[163] State agencies are required to comply with the recommendation issued by the HTCC.[164]

HTCC's review of the effectiveness, safety, and cost of CT colonography determined that the technology is not a covered benefit for routine colorectal cancer screening.[165]

Effectiveness

The HTCC report set forth that effectiveness was an area of substantial discussion. In particular they noted that radiologists require special training to use the technology. There was concern that clinical results may not be replicable in the community based setting in terms of the availability of equipment and ability to enforce training levels. Further, sensitivity evidence was mixed based on polyp size.[166]

Safety

HTCC was concerned about both radiation exposure from the technology and overall lifetime radiation accumulation from other tests. However, ultimately the HTCC found that CT colonography was either equivalent or safer than optical colonoscopy.[167]

Cost

HTCC was concerned over the cost of CT colonography. They felt that if the cost was reduced to one-third of optical colonoscopy then the technology might be more

---

[163] *Ibid.*

[164] *Ibid.*

[165] *Ibid*, page 1.

[166] *Ibid*, page 2.

[167] *Ibid*.

feasible from a cost perspective. Ultimately HTCC determined there was insufficient evidence to make a firm determination about cost.[168]

U.S. Preventative Services Task Force

The USPSTF provides "recommendations about preventive care services for patients without recognized signs or symptoms of the target condition. It bases its recommendations on a systematic review of the evidence of the benefits and harms and an assessment of the net benefit of the service."[169] USPSTF acknowledges that given their recommendation, it is still up to the physician to make a clinically appropriate decision regarding whether or not to use a particular technology based upon their assessment of an individual patient.[170]

USPSTF employs a grading system to evaluate new technologies. USPSTF's grades and suggestions for practice are set forth in Table 16.

[168] *Ibid.*

[169] U.S. Preventative Services Task Force. (2008). Screening for colorectal cancer: U.S. Preventative Services Task Force Recommendation. *Ann Intern Med*, 149:627-37, at 627.

[170] *Ibid.*

Table 16. USPSTF Grades and Suggestions for Practice[171]

Grade	Definition	Suggestion for Practice
A	USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer/provide this service.
B	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	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 other considerations support offering or providing the service in an individual patient.
D	USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefit.	Discourage the use of this service.
I Statement	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 the service is offered, patients should understand the uncertainty about the balance of benefits and harms.

USPSTF's levels of certainty regarding the net benefit of a technology are set forth in Table 17.

Table 17. USPSTF Levels of Certainty Regarding Net Benefit[172]

Level of Certainty	Description
High	The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.
Moderate	The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: • The number, size or quality of individual studies • Inconsistency of findings across individual studies • Limited generalizability of findings to routine primary care practice • Lack of coherence in the chain of evidence As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion.
Low	The available evidence is insufficient to assess effects on health outcomes. Evidence is insufficient because of: • Limited number or size of studies • Important flaws in study design or methods • Inconsistency of findings across individual studies • Gaps in the chain of evidence • Findings that are not generalizable to routine primary care practice • A lack of information on important health outcomes More information may allow an estimation of effects on health outcomes.

[171] *Ibid*, at 636.

[172] *Ibid.*

USPSTF reviewed CT colonography in 2008. In their analysis, they concluded that there was insufficient evidence to assess the harms related to extracolonic findings.[173] This finding was based on four considerations: (1) USPSTF found that CT colonography could help reduce cancer mortality if patients would submit to this form of testing and not other forms; however, (2) the potential harms of CT colonography could be significant, particularly regarding lifetime cumulative radiation risk from the test as well as other tests that involve radiation exposure; further, (3) radiologists need specialized training in order to be proficient with the test; and (4) the overall cost of CT colonography is high because bowel preparation is still required.[174] As a result, USPSTF recommends randomized clinical trials to clearly define the benefits and harms of the technology.[175]

Discussion of the Different Assessments and Guidances

Of the five HTA organizations that issued assessments and guidance regarding CT colonography, two approved the technology for use and three did not. Each HTA organization has a different mission and purpose. As a result, it is not surprising that there are different conclusions among HTA organizations regarding the same technology. Thus, the issue of consistency is not necessarily among the separate findings of each HTA organization. However, consistency does become an issue when the technology assessment is counter to the stated objectives and/or definitions of desirable types of evidence for the HTA organization. Table 18 summarizes the key criteria across all 5 HTA organizations regarding CT colonography.

---

[173] *Ibid*, at 629.

[174] *Ibid,* at 630.

[175] *Ibid*, at 631.

Table 18. Summary of Different HTA Criteria for CT Colonography

Criteria	CTAF	TEC	NICE	HTCC	USPSTF
Regulatory Approval	✓	✓	✓	✓	✓
Evidence allows conclusions about health outcomes	✓	✓	✓	X	✓
Shows improvement of health outcomes	X	✓	✓	X	✓
As beneficial as any established alternatives	X	✓	✓	X	X
Improvement attainable outside investigational setting	X	✓	✓	X	X
Cost	N/A	N/A	✓	X	X

CTAF and TEC are two organizations with almost identical HTA criteria. Both organizations agree that the technology meets the first two criteria: (1) approval from the appropriate government body; and (2) scientific evidence permits conclusions concerning the effectiveness of the technology.

The third criterion is the point of divergence between CTAF and TEC: improving the net health outcome. CTAF argues that the potential harms of the technology outweigh the benefits. Specifically, follow up regarding extracolonic findings is a key area of difficulty for CTAF. TEC, on the other hand, found an inference of effectiveness that allowed the panel to issue a positive finding for this criterion.

The fourth criterion is also a point of divergence. CTAF issued concerns related to radiation exposure and the identification of extra-colonic findings. While TEC found that the benefit of identifying and removing cancer outweighed the small, unquantifiable, risks of detecting unrelated health problems and radiation exposure. Lastly, the fifth criterion is also different with CTAF finding CT colonography to be too difficult to offer in community-based settings while TEC held that the ACRIN data demonstrated that radiologists could be trained to use the technology effectively.

The disagreement between CTAF and TEC outlines a question of level of evidence required. CTAF expects a higher level of evidence than TEC. TEC supports the data set forth in ACRIN while CTAF acknowledges but does not accept the ACRIN findings as sufficient to sway their decision.

NICE, HTCC and USPSTF examined different criteria than CTAF and TEC. The guidance issued by NICE supports the use of the technology based on sensitivity and specificity data. HTCC based its review on the effectiveness, safety, and cost of the technology that was insufficient for a positive finding. Lastly, USPSTF based its review on benefits and harms of the technology, and included comments about costs. In the case of CT colonography, the harms outweighed the benefits.

Conclusion

The area of commonality among all of these health technology assessments of CT colonography is that the levels of evidence used by the agencies are inconsistent. These inconsistencies become particularly apparent when reviewing a screening technology to be used in a patient population that is asymptomatic because evidence of benefits and harms become more subjective. While understanding whether a screening technology has achieved the requisite approval from the governmental regulatory authority is relatively straightforward, establishing the efficacy, effect on health outcomes, and whether or not the technology is as beneficial as established alternatives are not as evident. Differences of HTA agency missions and funding are a likely reason for the inconsistencies at one level. However, whether or not a more robust set of criteria could be used in assessing screening technologies is a question that may be further investigated via interviews with HTA decision makers.

Chapter 5: Interview Results

Interviews

Purposeful sampling was used to select respondents to participate in this study. First, HTA organizations were selected for inclusion in the sample based on the literature review. Next, leaders within those organizations were researched. Last, Clifford Goodman, PhD, health technology assessment expert at The Lewin Group was consulted regarding the list of organizations and corresponding leaders.[176] Dr. Goodman provided feedback to focus the list of respondents on exceptional leaders in the field of health technology assessment in the United States, Canada, Germany, and the United Kingdom.

A total of 20 individuals from 16 different technology assessment organizations and 4 countries were contacted to participate in the study. Respondents were selected for participation on the basis of their responsibilities related to the implementation of health technology assessment programs within their organizations, or their participation in the creation of assessments. Individuals were recruited via email. As a part of the invitation, respondents were informed that the purpose of the interview was to develop a best practice for health technology assessment of screening technologies, see Appendix 1. Recruitment continued until 12 respondents agreed to participate in the study.

---

[176] The Lewin Group is a nationally recognized healthcare consulting firm with consultants drawn from industry, government, academia. Many of The Lewin Group's consultants are national authorities whose strategies for health and human services system improvements come from their personal experience with imperatives for change. The Lewin Group is often consulted by government agencies, including the office of the President of the United States, to provide guidance regarding healthcare matters.

Twelve of the 20 (60%) individuals invited participated in the interviews. Of the 8 individuals that were not included in the study, 4 (50%) were unable to participate, citing busy work schedules. The remaining 4 (50%) did not respond to the email or two follow up recruitment emails.

The 12 respondents represent 11 different health technology assessment organizations in 3 countries: United States, Canada and United Kingdom. While healthcare in Canada and the United Kingdom are delivered via national mechanisms, like the United States, both countries have employed HTA as a method for making decisions about accessibility to technologies. Table 19 and Figure 2, respectively, provide an overview of the distribution of participating organizations across the types of health technology assessment organization they represent. Table 20 provides an overview of the distribution of non-participating organizations across the same variable as Table 19.

Table 19. Distribution of Participating Organizations

Respondent	Type of HTA Organization	Source of Funding	Findings Directly Determine Coverage for Health Plan (Y/N)
Respondent 1	For Profit Organization	Private Contracts	N
Respondent 2	Government Agency	Government	Y
Respondent 3	Government Agency	Government	Y
Respondent 4	Not for Profit Organization	Government and Donation	Y
Respondent 5	Not for Profit Organization	Government and Donation	Y
Respondent 6	Academic Medical Center	Multiple Funding Sources	N
Respondent 7	Academic Medical Center	Multiple Funding Sources	N
Respondent 8	Government Agency	Government	Y
Respondent 9	Government Agency	Government	Y
Respondent 10	Private Payer	Private Contracts	Y
Respondent 11	Not for Profit Organization	Multiple Funding Sources	N
Respondent 12	Government Agency	Government	Y

Figure 2. Distribution of Participating Organizations by Type of Organization

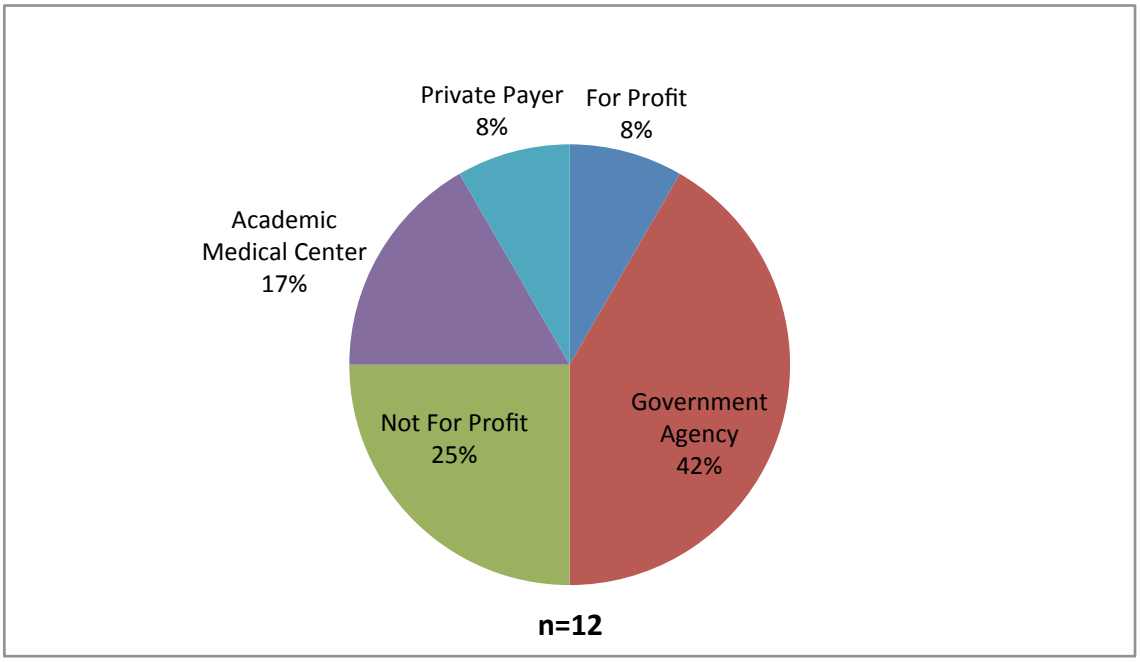


Table 20. Distribution of Non-Participating Organizations

Respondent	Type of Organization	Source of Funding	Findings Directly Determine Coverage for Health Plan (Y/N)
Respondent 1	Academic Medical Center	Multiple Funding Sources	N
Respondent 2	Not for Profit Organization	Donation	N
Respondent 3	Government Agency	Government	Y
Respondent 4	Not for Profit Organization	Multiple Funding Sources	N
Respondent 5	Not for Profit Organization	Multiple Funding Sources	N
Respondent 6	Not for Profit Organization	Multiple Funding Sources	Y
Respondent 7	Government Agency	Government	Y
Respondent 8	Government Agency	Government	Y

Interviews were conducted individually with each respondent. Ten of the 12 interviews were with leaders within their organization. In a health technology assessment organization the titles of the leaders include: president, executive director, director, and program manager. The remaining 2 respondents were research analysts within their

organizations. These individuals were recommended by their respective leaders to participate in the interviews because they were the most knowledgeable people about specific health technology assessment criteria.

Interviews took place from in May and June 2011. Due to the geographic location of respondents and the cost parameters of this study, interviews were conducted via telephone. Respondents were informed that the information collected during the interview would be kept confidential. Specific answers would not be attributed to the respondent or their organization. The interview results would only be used in summary form. The interview guide and confidentiality terms are in Appendix 2.

The sources of the questions posed were from the literature and discussions with researchers who have previously studied the process and use of HTA evaluations. The interview guide was pilot tested via cognitive interviews with 3 individuals who were knowledgeable about health technology assessment. The interviews with respondents were semi-structured. Modifications were made during the interview based on the respondent's responses to the preceding questions. Interview telephone calls were not recorded. To ensure accuracy, the interviewer and one research assistant took handwritten notes to record responses. Respondents were provided the option to review and edit the notes from the interview at their request.

Data Analysis

The information obtained from the interviews was both quantitative and qualitative. The quantitative portion of the interview involved several questions with respondents providing a numeric rating. The rating was based upon a five point Likert scale in which 1 was "strongly disagree", 2 was "disagree", 3 was "neither agree nor

disagree", 4 was "agree", and 5 was "strongly agree". In this section, the figures display the respondent's answers to questions asked in the interview guide. Blue circles represent respondents based in the United States and red circles represent respondents based outside of the United States. Where the data had sufficient differentiation, the center line of "box and whisker" diagrams represents the median. The green shaded box encloses the 25 and 75 percentiles. The extended lines, or "whiskers", represent the minimum and maximum of the respondents' answers.

The majority of the interview information was qualitative. The analysis of the qualitative information involved a review of the interview transcripts to identify themes and to compare and contrast responses across interviews. The volume of interview data was minimal. As a result, themes were manually coded, and where possible and appropriate, counted and weighted by frequency of mention. Themes were grouped for discussion and conclusion purposes.

Descriptive Analysis

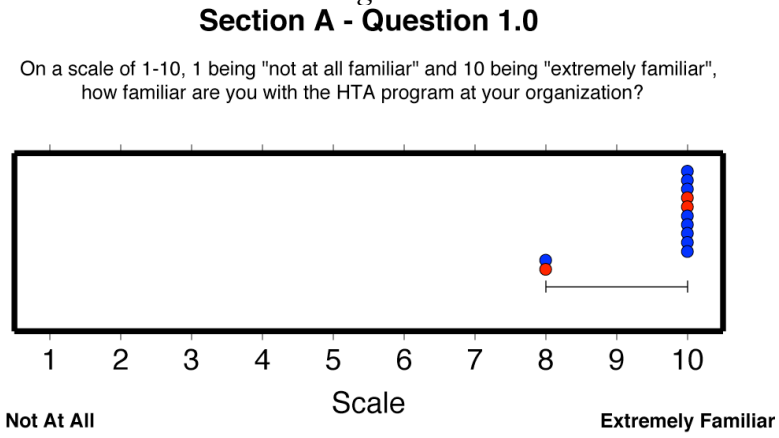
The interview was divided into three sections. All 12 respondents answered the questions in each section. Some respondents elected to not answer some specific questions and these are specifically noted. The three sections of the interview included: (1) Health Technology Assessment at Your Organization; (2) Value and Obtainability of Outcomes Evidence; and (3) Evaluation of Cost-Effectiveness of New Screening Technologies.

1. Health Technology Assessment at Your Organization

Questions in the first section of the interview were quantitative and addressed details about health technology assessment at the respondent's organization. In terms of

familiarity with the health technology assessment program at their organization, 10 (83%) out of 12 respondents represented their knowledge of HTA as a 10 on a scale of 1 to 10 in which 1 is not at all familiar and 10 is extremely familiar. The remaining 2 (17%) respondents represented their level of knowledge as an 8 on the 10 point scale. As a result, the 12 respondents were all very familiar with their HTA programs.

Figure 3.



There are a number of critical components for the development of HTA programs at organizations. Respondents were asked a series of questions to understand the most important components for development of HTA programs at their organizations.

Seven (58%) out of 12 respondents strongly agreed that HTA is a part of the mission statement of their organization. Organizations represented included: 1 for profit company (n=1), 2 government agencies (n=5), 1 academic medical center (n=2), 2 not for profit organizations (n=3), and 1 private payer (n=1). One of the 6 respondents indicated:

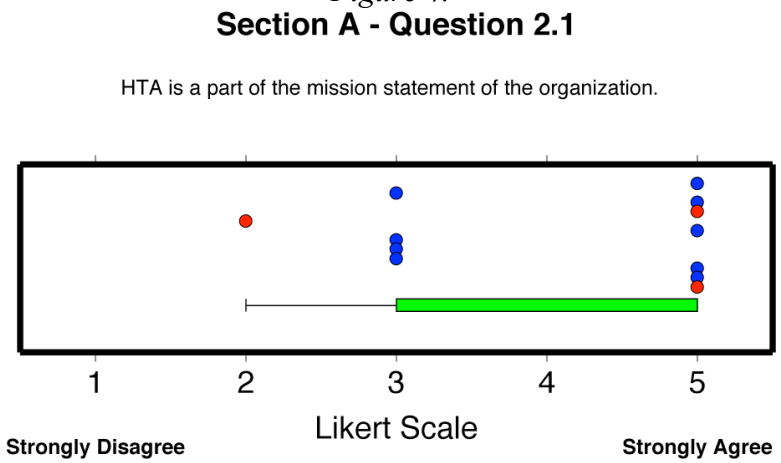
> We don't have a mission statement. If the word "mission" wasn't in there, then the statement would be entirely accurate, but I will strongly agree since it is what we do.

Four (33%) of the respondents neither agreed nor disagreed that HTA is a part of the mission statement of the organization. These organizations represented 3 government

agencies (n=5) and 1 academic medical center (n=2). One of the 4 respondents commented: "HTA is implicit in our mission statement, not explicit."

One (8%) respondent that disagreed that HTA was a part of the mission statement is a part of a large not for profit (n=3) that has a diverse mission for a local government.

Figure 4.

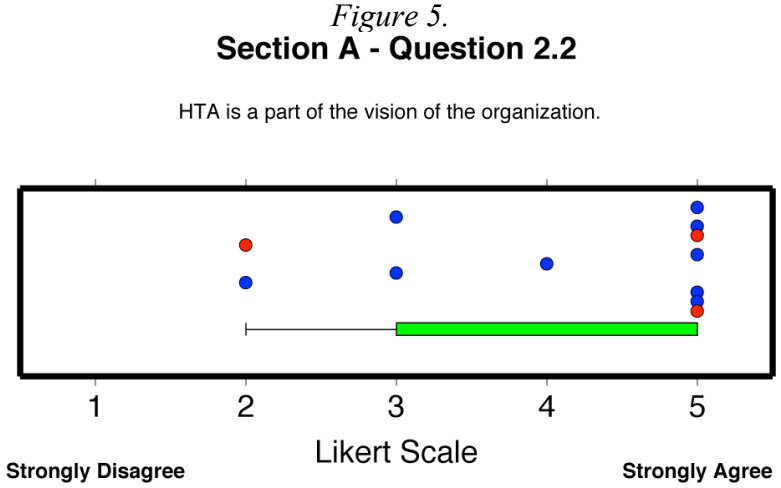


Seven (58%) out of 12 respondents indicated that HTA is a part of the vision of the organization. Organizations represented included: 1 not for profit (n=1), 2 government agencies (n=5), 2 not for profit (n=3), and 1 academic medical center (n=2), 1 private payer (n=1).

One (8%) academic medical center (n=2) agreed.

Two respondents (17%) disagreed representing a not for profit (n=3) and a government agency (n=5). Two respondents (16%) neither agreed nor disagreed. Both of these respondents represented government agencies (n=5).

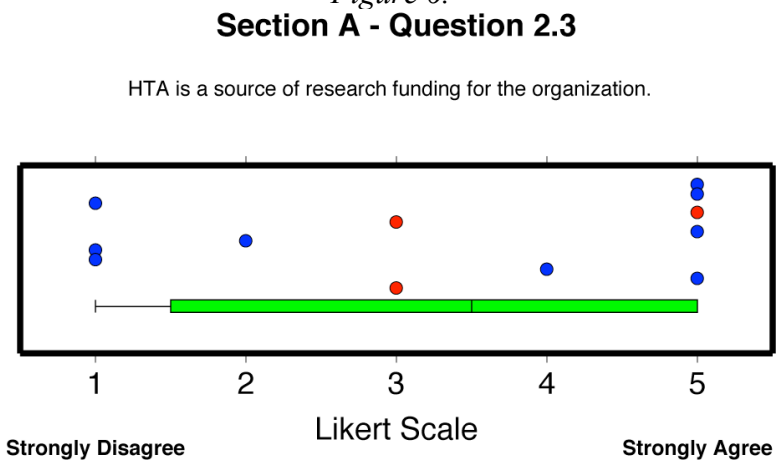
Figure 5.



Five (42%) respondents strongly agreed that HTA is a source of research funding for the organization. Types of organizations included: 1 for profit (n=1), 1 government agency (n=5), 2 not for profit (n=3), and 1 academic medical center (n=2).

One (8%) private payer (n=1) agreed. Two respondents (17%), 1 not for profit (n=3) and 1 government agency (n=5), neither agreed nor disagreed. One (8%) academic medical organization (n=2) disagreed. Three (25%) government agency organizations (n=5) strongly disagreed.

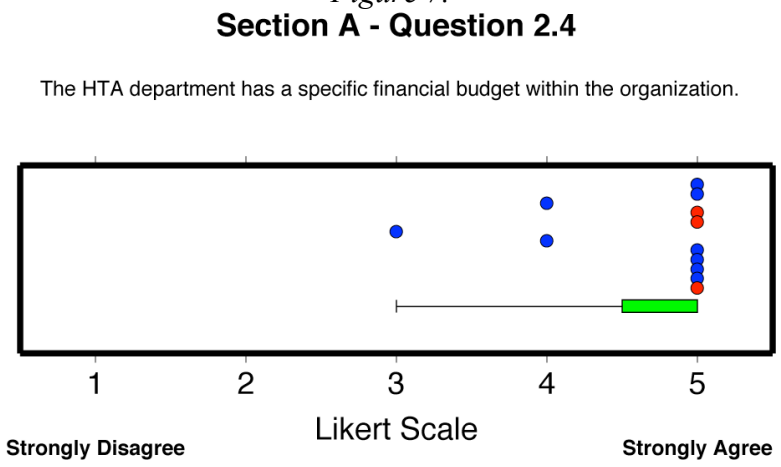
Figure 6.



Nine (75%) respondents strongly agreed that the HTA department has a specific financial budget within the organization. Types of organizations included: 1 for profit (n=1), 4 government agencies (n=5), 3 not for profit (n=3), and 1 private payer (n=1).

Two (17%) respondents, representing 1 government agency (n=5) and 1 academic medical center (n=2) agreed. One (8%) academic medical center (n=2) neither agreed nor disagreed.

Figure 7.



No respondents strongly agreed that HTA is the sole purpose of their organization. Two (17%) respondents agreed. Organizations represented included: 1 academic medical center (n=2) and 1 private payer (n=1).

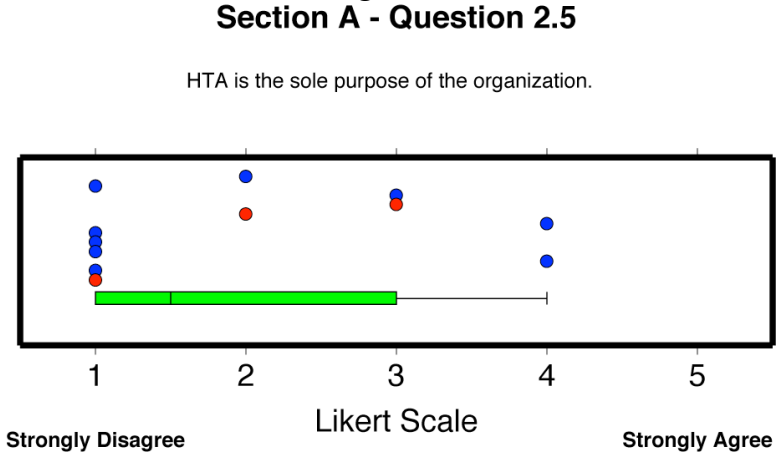
Two (17%) respondents neither agreed nor disagreed. Organizations represented included: 1 government agency (n=5) and 1 not for profit (n=3).

Two (17%) respondents disagreed. Organizations represented included: 1 for profit (n=1) and 1 not for profit (n=3). One of the respondents commented to clarify the response:

> Our core purpose is HTAs, not sole or only, so I would have to disagree with your statement.

Six (50%) respondents strongly disagreed. Organizations represented included: 4 government agencies (n=5), 1 not for profit (n=3), and 1 academic medical center (n=2).

Figure 8.



One (8%) respondent, not for profit (n=3), strongly agreed that HTA drives coverage decisions at their organization.

Two (17%) respondents, 1 academic medical center (n=2) and 1 government agency (n=5), agreed. One (8%) respondent, private payer (n=1), neither agreed nor disagreed. Two (17%) respondents, for profit (n=1) and government agency (n=5), disagreed. One respondent commented:

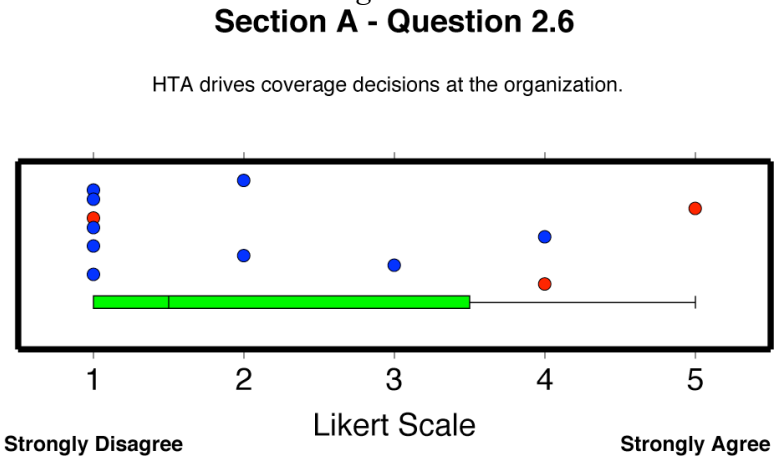
> We don't make coverage decisions. We provide content to drive coverage decisions at other organizations, our clients. HTA is a core component to drive decision making at other organizations, like health plans, government agencies, providers.

Six (50%) respondents strongly disagreed. Organizations represented included: 4 government agencies (n=5), 1 academic medical center (n=2), and 1 not for profit (n=3). One respondent commented on their function in the process:

> We don't actually make decisions here. It drives coverage decisions elsewhere, but not here. We exist to do the assessments, and provide conclusions and/or advice, and

> the decisions are then made...We are at arms length for the decisions, but we are involved in the assessment process.

Figure 9.



The quantitative portion of this section concluded and a qualitative section followed. All 12 respondents were asked if their organization has stated goals and objectives for their HTA program. Ten (83%) respondents indicated that their organization has stated goals or objectives for their HTA program. Two of these respondents specifically commented on making decisions that are evidence-based when describing their organization's goals and objectives. One respondent linked evidence-based HTAs to a separate organizational goal of making coverage decisions:

> Our goals are really to make evidence-based coverage decisions. But a subsidiary goal is to base those goals, those evidence-based decisions, on structured reviews of evidence, i.e., an HTA. Our goal isn't to produce HTAs for the sake of HTA, since we are both the producer and the end user.

Another respondent indicated that their organizational goal is not to make decisions, rather to weigh "the valance of benefits, harm and costs of adopting or covering new technologies." Similarly, a separate respondent spoke to the organizational goal to "rigorously summarize, evaluate and communicate unbiased information."

Two (17%) respondents indicated that their organization does not have stated goals or objectives for their HTA program. One of the respondents indicated that their organization does HTA work on a contract basis, as a result, there is not an organizational stated goal or objective related to HTA. The other respondent indicated that their organization is going through a lot of change and they are revising their organizational goals and objectives. Table 21 summarizes the responses by type of organization.

Table 21. Stated Goals and Objectives for HTA by Organization Type

Organization Type	Goals & Objectives	No Goals & Objectives
For Profit	√	
Government Agency	√√√√√	
Not for Profit	√√	√
Academic Medical Center	√	√
Private Payer	√	
Total	**10**	**2**

All 12 respondents were asked when the HTA program was implemented at their organization. Ten (83%) respondents indicated that the HTA program was implemented prior to 2000 and 2 (17%) indicated the HTA program was implemented after 2000. Table 22 summarizes the responses by type of organization.

Table 22. Timing of HTA Program Implementation by Organization Type

Organization Type	Before 2000	After 2000
For Profit	√	
Government Agency	√√√√√	
Not for Profit	√√	√
Academic Medical Center	√	√
Private Payer	√	
Total	**10**	**2**

All 12 respondents were asked how many HTAs their organization had conducted since 2000. Seven (58%) respondents indicated that their organization has conducted between 30 and 40 HTAs since 2000. Five (42%) respondents indicated that their

organization has conducted over 40 HTAs since 2000. Table 23 summarizes the responses by type of organization.

Table 23. Number of HTAs Conducted Since 2000 by Organization Type

Organization Type	30-40	>40
For Profit	√	
Government Agency	√√	√√√
Not for Profit	√√√	
Academic Medical Center	√	√
Private Payer		√
Total	**7**	**5**

All 12 respondents were asked to share the criteria for HTA at their organization. None of the respondents provided specific criteria; however, they did share the general type of criteria. One (8%) respondent, representing a for profit, indicated that their organization uses proprietary criteria:

> We have developed our own evidence scoring method…our criteria are an evidence score, which considers multiple aspects. It looks at research design, but goes beyond. [It] looks at implementation and execution of research design and methodology, research execution, the size of effect, the research statistics, research replication and generalizability. Evidence scores are specific to patient indications and patient selection criteria. A technology may have multiple grades or scores based on patient indication and use.

Seven (58%) respondents indicated that their criteria for HTA varied. The rationale for the variance included that different assessments would use different criteria, and that the assessment would vary based on the scope of information and analysis the client or agency funding the HTA was asking for in the HTA.

Two (17%) respondents referred to their HTAs as systematic reviews of evidence and literature. In addition, 2 (17%) respondents indicated that the HTA criterion were

driven by "issues of national importance." Table 24 sets forth the criteria for HTA by organization type.

Table 24. Criteria for HTA by Organization Type

Organization Type	Proprietary Criteria	Criteria Varies by Type of HTA	Systematic Review	National Importance
For Profit	√			
Government Agency		√√√		√√
Not for Profit		√√	√	
Academic Medical Center		√	√	
Private Payer		√		
Total	**1**	**7**	**2**	**2**

All 12 respondents were asked if their criteria and data requirements were published and available. Ten (83%) respondents indicated that the criteria and data requirements were published and available. Two (17%) respondents indicated that the criteria and data requirements were not available. One of the reasons provided for the lack of availability was not stated, while the other reason was due to changes in the organization. Table 25 sets forth the availability of HTA criteria by organization type.

Table 25. Availability of HTA Criteria by Organization Type

Organization Type	Available	Not Available
For Profit	√	
Government Agency	√√√√	
Not for Profit	√	√√
Academic Medical Center	√√	
Private Payer	√	
Total	**10**	**2**

All 12 respondents were asked how their organizations select technologies to assess. Eight (67%) respondents indicated that technologies are selected based on client request. In this context, client is defined as the entity for which the respondent's organization is conducting the HTA. For example, a larger government agency, healthcare systems, managed care organizations, and Ministries of Health. While it is

possible that a client could be a patient group or defined group of health consumers, no respondent provided these groups as examples of clients. Several respondents commented that technologies are selected if there is a "controversy" about the use of the technology that may be triggered by "demands or pressures from specialty groups." Specialty groups may be patients, patient advocacy groups, manufacturers of screening technologies, physicians, and payers. Three (25%) respondents indicated that technologies are selected if they rise to the level of national importance. One respondent commented:

> If there is significant disagreement, then that's the entry criteria. Not that we have the infrastructure of budget to look at all the criteria, but that gets you through the door.

One (8%) respondent indicated that products are selected based upon selection criteria. Examples of these criteria include the burden of disease in terms of population affected, morbidity and mortality; resource impact in terms of the cost to the health system; policy importance in terms of whether the topic falls with a government priority area; whether there is inappropriate variation in practice across the country; and whether there are factors affecting the timeliness or urgency for a guidance to be produced. Table 26 sets forth the selection of technologies for HTA by organization type.

Table 26. Selection of Technologies for HTA by Organization Type

Organization Type	Client-Driven	National Importance	Selection Criteria
For Profit	√		
Government Agency	√	√√√	√
Not for Profit	√√√		
Academic Medical Center	√√		
Private Payer	√		
Total	**8**	**3**	**1**

All 12 respondents were asked to describe an HTA decision that was made at their organization. All 12 (100%) respondents indicated that their organizations do not make HTA decisions, rather, they "synthesize and summarize evidence." In addition, all 12 (100%) respondents indicated that an existing HTA might change in the future as more evidence for the technology becomes available. One respondent provided an example of actively looking for new information by "conducting scans and constant literature reviews for new information."

All 12 (100%) respondents use a committee or team approach to conduct the HTA. Teams are comprised of clinical experts, medical librarians, and researchers. Two respondents also included economists and members of the community. One respondent included product manufacturers. Assessments are either delivered directly to clients or are posted on websites.

None of the respondents had explicit metrics used by their organization to measure the effectiveness of a particular HTA. One respondent commented that they "use productivity and impact measures" but these measures were not directly linked to the overall efficacy of the HTA. Another respondent commented:

> We do try, but it is difficult to measure at a systems level. There can be confounding, and can be difficult to attribute to [our organization], particularly around outcomes and budget impact. We do know our impact in diseases with independent technologies and we can see [impact] to the uptake of technologies based on a yes or no [HTA]…overall it is very difficult to have generic metrics measure this type of impact.

Ten (83%) respondents indicated that health policy did change once the assessment was complete. One respondent provided an example:

> One of the most interesting ones, because we tracked it for so long, would be PET scans. That technology was considered by most to be investigational, not ready for clinical application when it was first introduced about 25 years ago. It has steadily evolved in context of use, and value demonstrated. Initially, there was a lot of push back from a coverage point of view, because the data was so limited at the time of its introduction on its value and utility. Over the 20 year period we have tracked PET scans, it has experienced multiple avenues for coverage and reimbursement positive decisions. So, that has changed from limited coverage to wide spread coverage. When a technology is first introduced, there is usually not a lot of evidence, the utility isn't clear, and the patient selection criteria are unclear. As the body of evidence improves, our understanding of its use and value improves. New technology usually would starts at a [low] evidence score or grade, and then moves up from there as data and evidence are generated in the marketplace

One (8%) respondent did not provide a comment and the other respondent (8%) indicated that their organization was not tracking if policy changed over time. The organizations represented were both government agencies.

All 12 respondents were asked if feedback from providers, patients or manufacturers caused the organization to change or re-evaluate the HTA. Seven (58%) respondents indicated that feedback might cause a change or re-evaluation of an HTA. One respondent indicated that while feedback has been received and considered, to date, no HTA has required change. Another respondent commented that their organization is "required by law to post a decision and accept 30 days of public comment on it." In this respondent's case, while feedback may come from any one in the public, they specifically look "for feedback that speaks to medical evidence." Another respondent indicated that they receive feedback through several mechanisms. They receive information via an

appeal process once an HTA is published, through a consultation process with the public, judicial review, and through a process of "interrogating physician and patient experts.

Three (25%) respondents indicated that their organization does not collect or consider feedback. Organizations included: 2 government agencies (n=5) and 1 academic medical center (n=2). Two (17%) respondents, government agency and academic medical center, did not provide comment.

All 12 (100%) respondents confirmed that they do have an ongoing or multi-year budget commitment for HTA at their organization. The duration of the budget commitment ranged from 1 year to 5 years. Table 27 sets for the duration of budget commitment by organization type.

Table 27. Duration of Budget Commitment by Organization Type

Organization Type	1 year	2-3 years	5 years	Multi-Year Unspecified
For Profit				√
Government Agency	√√√√			√
Not for Profit			√√	√
Academic Medical Center			√√	
Private Payer		√		
Total	**4**	**1**	**4**	**3**

All 12 respondents were asked if they knew the average cost of conducting an HTA at their organization. Four (33%) respondents indicated that the cost of the assessment depended upon the complexity of the HTA. As a result, a specific estimate could not be provided. One respondent commented:

> We prepare a lot of report types, from simple to sophisticated. With such a wide range, an average would not be very helpful…The custom type responses would be several hundred thousand dollars, unlike the simple, quick turnaround assessment. The large assessments can take more than a year to complete. They are very resource intensive.

Seven (58%) respondents provided estimates ranging from $15,000 to $300,000. One (8%) respondent did not have sufficient information to comment. Table 28 sets forth the ranges of estimates based on organization type for the 6 respondents that provided information.

Table 28. HTA Budget Estimate by Organization Type

Organization Type	$15,000	$150,000-$200,000	$300,000
For Profit			
Government Agency	√√	√√	
Not for Profit		√	√
Academic Medical Center		√	
Private Payer			
Total	**2**	**4**	**1**

All 12 respondents were asked if they felt there were specific issues that organizations seeking to conduct HTA of screening products should know either before they begin, during the analysis, or when the analysis is complete. All 12 respondents identified specific issues that should be considered. Interestingly, no respondent mentioned the cost or burden from a false positive from a test as an issue. These issues are summarized in Table 29.

Table 29. Summary of Issues to Consider in Conducting an HTA for a Screening Product

Issue	Issues Identified by Organizations				
	FP	GA	NFP	AMC	PP
Reliability, Sensitivity, and Specificity	√	√	√	√	√
Identification of Appropriate Patient Population	√	√			
Link Between Test and Actual Expression of Disease	√	√			
Availability of Treatment Options	√		√		
Patient Anxiety	√		√		
Length of Time To Gather Sufficient Data of Efficacy	√				
Natural Course of Disease		√			√
Ethical Considerations			√	√	√
Availability of Treatment Options if Disease is Identified	√	√			
Cost-Effectiveness and Cost of Alternatives		√			
Patient Acceptance of Test			√		

Key: FP = For Profit GA = Government Agency
NFP = Not for Profit AMC = Academic Medical Center
PP = Private Payer

One of the respondents concluded with the following statement:

> Screening healthy people has a high ethical obligation to do no harm. Every intervention has that and screening magnifies that because you're actively intervening with an otherwise healthy population. It's a more proactive framework than treating those that symptomatically present themselves for treatment.

Another respondent encouraged: "keep things open for discussion, allow some degree of contestability, manage vested interests from where these come from."

2. Value and Attainability of Outcomes Evidence

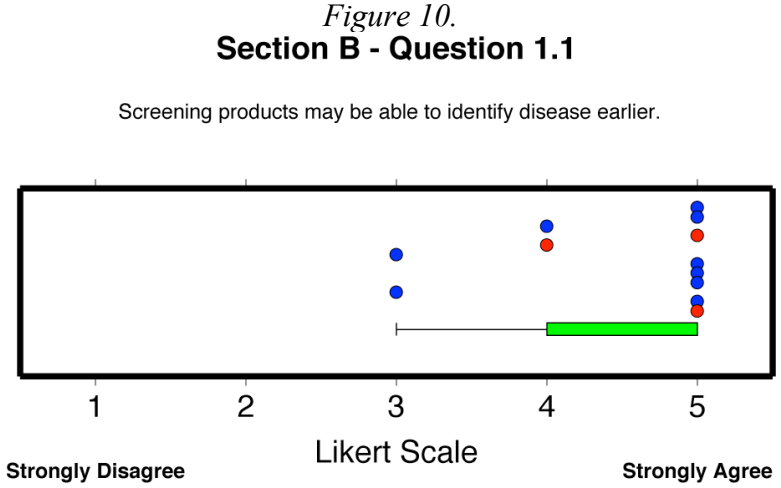
Screening technologies may be able to help identify whether or not a patient has a specific disease in order to inform treatment decisions before symptoms occur. Controversies about screening technologies exist. In this quantitative section of the interview, respondents were asked to respond using a five point Likert scale to indicate whether they agree or disagree with a particular screening product controversy.

Eight (67%) respondents strongly agreed that screening products may be able to identify disease earlier. Organizations represented included: 1 for profit (n=1), 4 government agency (n=5), 2 not for profit (n=3), and 1 academic medical center (n=2).

Two (17%) respondents agreed. Organizations represented included: 1 government agency (n=5) and 1 not for profit (n=3).

Two (17%) respondents neither agreed nor disagreed. Organizations represented included: 1 academic medical center (n=2) and 1 private payer (n=1).

Figure 10.

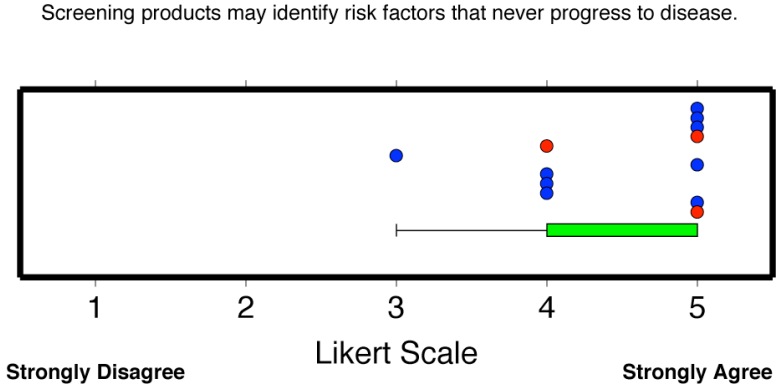


Seven (58%) respondents strongly agreed that screening products may identify risk factors that never progress to disease. Organizations represented include: 1 for profit (n=1), 3 government agencies (n=5), 2 not for profit (n=3), and 1 academic medical center (n=1).

Four (33%) respondents agreed. Organizations represented included: 1 not for profit (n=3), 2 government agencies (n=5), and 1 private payer (n=1).

One (8%) respondent representing an academic medical center (n=2) neither agreed nor disagreed.

Figure 11.

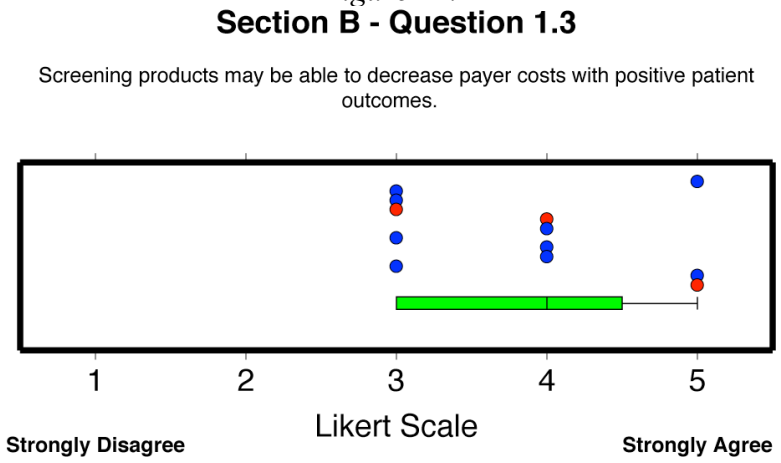


Three (25%) respondents strongly agreed that screening products may be able to decrease payer costs with positive patient outcomes. Organizations represented included: 1 for profit (n=1), 1 not for profit (n=3), and 1 government agency (n=5).

Four (33%) respondents agreed. Organizations represented included: 1 not for profit (n=3), 1 academic medical center (n=2), and 2 government agencies (n=5).

Five (42%) respondents neither agreed nor disagreed. Organizations represented included: 2 government agency (n=5), 1 not for profit (n=3), 1 academic medical center (n=2), and 1 private payer (n=1).

Figure 12.

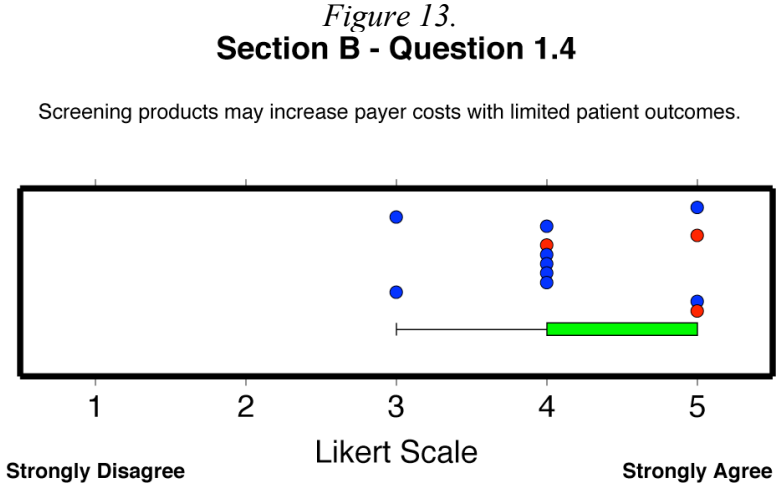


Four (33%) respondents strongly agreed that screening products may increase payer costs with limited patient outcomes. Organizations represented included: 1 for profit (n=1), 2 not for profit (n=3), and 1 government agency (n=5).

Six (50%) of respondents agreed. Organizations represented included: 3 government agencies (n=5), 1 not for profit (n=3), and 2 academic medical centers (n=2).

Two (17%) respondents neither agreed nor disagreed. Organizations represented included 1 government agency (n=5), and 1 private payer (n=1).

Figure 13.

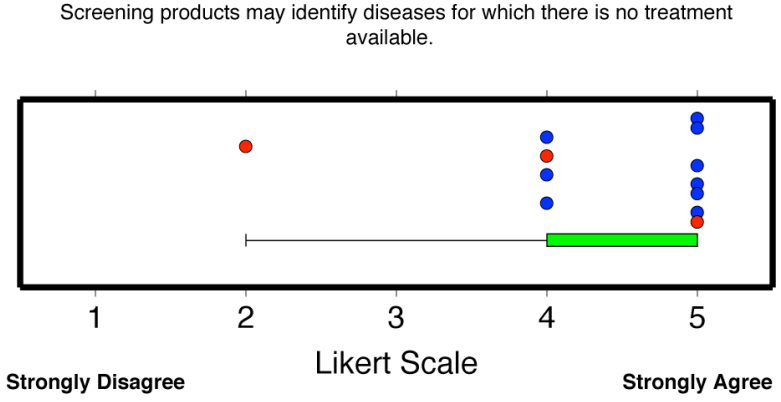


Seven (58%) respondents strongly agreed that screening products may identify diseases for which there is not treatment available. Organizations represented included: 1 for profit (n=1), 4 government agencies (n=5), 1 academic medical center (n=2), and 1 not for profit (n=3). One of the respondents commented that a screening product for Alzheimer's disease might be a relevant example.

Four (33%) respondents agreed. Organizations represented included: 1 government agency (n=5), 1 not for profit (n=3), 1 academic medical center (n=2), and 1 private payer (n=1).

One (<1%) respondent, representing a not for profit (n=3), disagreed.

Figure 14.



The quantitative portion of this section of the interview concluded and a qualitative section followed. All 12 respondents were asked if HTA criteria are different between therapeutic products and screening products. Eight (67%) respondents indicated that the criteria are different.

Four (33%) respondents indicated that the criteria are not different. These respondents indicated that while the criteria are not different, the way the analysis and type of evidence reviewed might be different. Specifically, one respondent commented: "[t]he process and how you might handle the evidence might differ since you might have to deal with more indirect evidence." Table 30 sets forth whether the HTA criteria for therapeutic and screening products are different by organization type.

Table 30. Are HTA Criteria Different Between Therapeutic and Screening Products by Organization Type

Organization Type	Criteria Are Different	Criteria Are Not Different
For Profit	√	
Government Agency	√√√√	√
Not for Profit	√	√√
Academic Medical Center	√√	
Private Payer		√
Total	**8**	**4**

All 12 respondents were asked what kinds of data are acceptable to support a HTA. Eight (67%) respondents indicated that it depends on the level of confidence that is needed. The higher the quality of the data in terms of the quality of the studies that have been done is important. One respondent commented:

> If data comes in from case studies and retrospective designs, then that weakens evidence. However, there are longitudinal cohort studies, like the Nurses Study and Framingham Health Study, which do provide insights in getting a handle on chronic conditions and treatments.

Three (25%) respondents indicated that they would take any kind of evidence. It did not have to be randomized clinical trial data. One respondent added that budget impact data and policy papers might also be relevant for an HTA. One (8%) respondent indicated that the data required between therapeutic and screening technologies were the same and that they would prefer randomized controlled trial data if it was available. Table 31 sets forth the responses by type of organization.

Table 31. Acceptable HTA data by Organization Type

Organization Type	Depends	Any Data	Randomized Control Trial
For Profit	√		
Government Agency	√√√√	√	
Not for Profit	√	√√	
Academic Medical Center	√		√
Private Payer	√		
Total	**8**	**3**	**1**

All 12 respondents were asked if there is a hierarchy of acceptable data, and if so, to share their hierarchy of data. Eight (66%) respondents indicated that there is a hierarchy of acceptable data. Of these 8 respondents, 5 (63%) indicated that randomized controlled trial was at the top of the data hierarchy. Organizations represented included: 4 government agencies (n=5), 3 not for profit (n=3), and 1 academic medical center (n=2). Two (25%) of the 8 respondents, not for profit (n=3) and academic medical center (n=2), did not provide a specific comment. One of the 8 respondents (8%), government agency (n=5), indicated: "the lowest point of entry would be single case studies or case series, observational or editorial."

Of the 12 respondents, 3 (25%) did not provide a comment regarding the presence of a hierarchy of data. Organizations represented included 1 for profit (n=1), 1 academic

medical center (n=2), and 1 private payer (n=1). One (8%) respondent, government agency (n=1) indicated that there is not a hierarchy of data.

3. Evaluation of Cost-Effectiveness of new Screening Technologies

Cost is not a specific criteria that is evaluated as a part of health technology assessments conducted in the United States. While some HTA organizations might consider cost, typically, cost is part of a separate study, if it is considered at all. However, some health technology assessments that are conducted outside of the United States take cost under consideration as a part of the HTA. In this quantitative section of the interview, respondents were asked to respond using a five point Likert scale to indicate whether they agree or disagree with statements related to HTA and cost.

Six (50%) respondents, 3 in the United States and 3 outside the United States, strongly agreed that cost should be added as an evidence measure in HTA. Organizations represented included: 1 for profit (n=1), 2 government agencies (n=5), 2 not for profit (n=3), and 1 academic medical center (n=2). One respondent located in the United States commented:

> It is foolish and not supportable for our society to avoid hard discussions around cost. Not sure if I would call it an evidence measure.

Another respondent located in the United States also commented on the accuracy of the statement "evidence measure" when referring to cost:

> Cost is not an evidence measure – it's an assessment measure. You use the evidence to build cost modeling and in the assessment. Be careful with the framework and language.

Two (17%) respondents agreed. Organizations represented included: 1 government agency (n=5) and 1 academic medical center (n=2). One respondent commented:

> Sometimes consideration of cost gets in the way of an assessment, but more often than not, it should be considered.

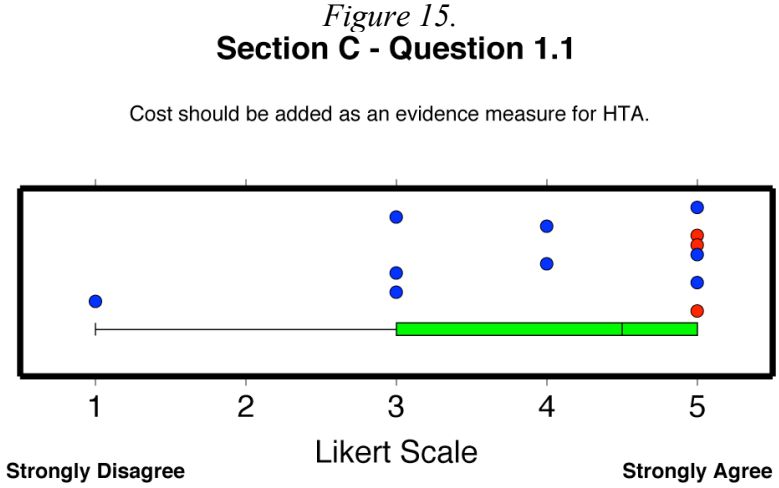
Three (25%) respondents neither agreed nor disagreed. Organizations represented included 2 government agencies (n=5) and 1 private payer (n=1). One respondent commented:

> Cost is defined as a rationing measure, rather than an evidence measure. If you have a finite budget, then it needs to be taken into account.

One (8%) respondent, representing a not for profit (n=3), strongly disagreed. This respondent commented that the role of the assessment at the organization was to set forth a statement of evidence based on facts. Cost does not have sufficient evidence to put forth an evidence-based conclusion:

> We have sections in our reports that speak to cost. In terms of cost-effectiveness, there is insufficient evidence of cost and cost-effectiveness to formulate generalizable evidence-based conclusions. We look at the pieces differently, and then we interpret. We put forth statements of evidence, but not recommendations.

Figure 15.



No respondents strongly agreed with the statement that cost is an unstated evidence measure for all HTAs. Two (17%) agreed that in general, cost is an unstated evidence measure for all HTAs. Organizations represented included 1 government agency (n=5) and 1 not for profit (n=3).

Three (25%) respondents neither agreed nor disagreed. Organizations represented included 2 government agencies (n=5) and 1 academic medical center (n=2). One respondent commented:

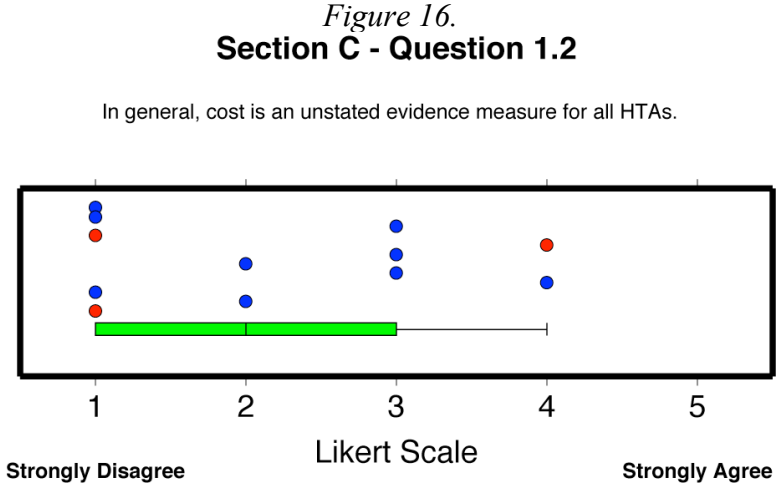
> Cost may have percolated at the top of the pile, but [it is] not taken into consideration as an evidence measure.

Two (17%) respondents, representing 1 academic medical center (n=2) and 1 not for profit (n=3) disagreed. One respondent commented:

> Every AMCP [Academy of Managed Care Pharmacy] dossier has a cost analysis in it that provides rationale for including it. [Cost is] only unusual if [there is a] political context. It is routine to have cost information.

Five (42%) respondents strongly disagreed. Organizations represented included: 1 for profit (n=1), 2 government agencies (n=5), 1 not for profit (n=3), and 1 private payer (n=1).

Figure 16.



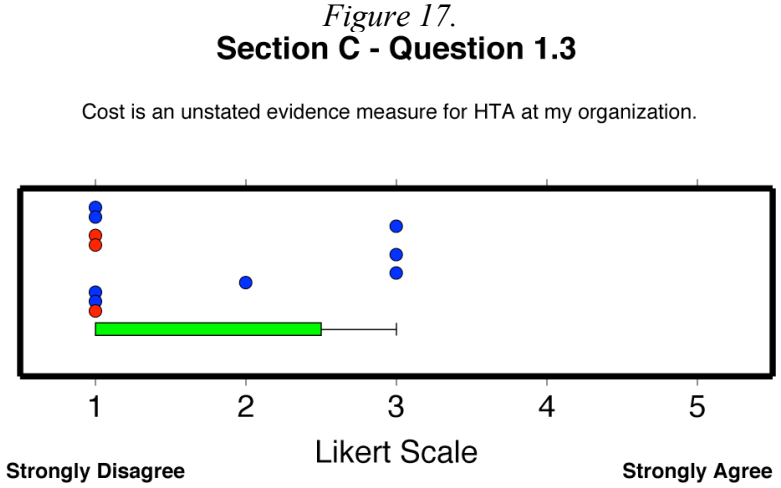
No respondents strongly agreed or agreed with the statement that cost is an unstated evidence measure for HTA at my organization. Three (25%) neither agreed nor disagreed. Organizations represented included: 2 government agencies (n=5) and 1 academic medical center (n=1). One respondent commented:

> Around colorectal cancer screening, Congress says we can consider cost, but generally, we don't.

Seven (58%) strongly disagreed. Organizations represented included: 1 for profit (n=1), 2 government agencies (n=5), 3 not for profit (n=3), and 1 private payer (n=1). One of the respondents located in the United States commented:

> Cost is an overt consideration at my organization. If we can get our hands on it, we know we will want to report it.

Figure 17.



One (8%) respondent, representing an academic medical center (n=2), strongly agreed with the statement that cost is not a relevant evidence measure for HTA. No respondents agreed.

Two (17%) respondents, representing a private payer (n=1) and not for profit (n=3), neither agreed nor disagreed.

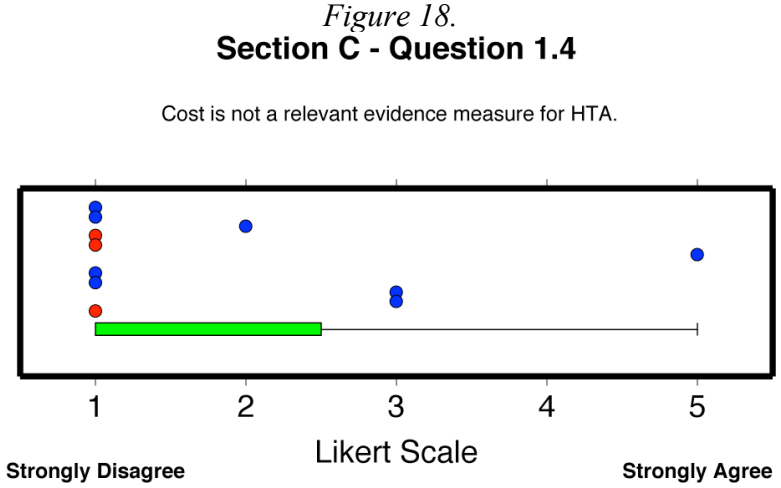
One (8%) respondent, representing a government agency (n=5), disagreed.

Seven (58%) respondents strongly disagreed. Organizations represented included: 1 for profit (n=1), 4 government agencies (n=5), and 2 not for profit (n=3). One respondent located in the United States provided additional clarification:

> It [cost] is relevant. [It] inform[s] policy development. There is a distinction between clinical evidence score and overall purpose of HTA. The HTA includes the clinical scoring, but also includes other elements important to decision makers. Systematic review evaluates body of evidence. HTA incorporates systematic review and evidence scoring, but goes beyond that so includes information [such as cost] for use in policy context.

One respondent, representing an academic medical center (n=2), declined to answer.

Figure 18.



Key findings for each of the three sections of the interview are set forth in the next chapter. These findings from the literature review and interviews form the basis for the creation of more robust criteria for screening products in the plan for change.

Chapter 6: Discussion of Interview Results

The objective of this research study is to evaluate existing technology assessment standards for screening technologies in order to establish a best practice that may be implemented by HTA organizations in the United States in the future. A literature review, scan of HTAs for screening products, and interviews with leaders in the field of HTA revealed that there is not a consistent approach to HTAs for screening technologies. This confirms that screening products offer a useful case study in where and how HTA as currently practiced yields differing findings by HTAs reviewing the same screening technology.

While screening products may have an impact on the care of patients in terms of identifying disease and appropriate treatment options at an early stage, confounders exist. For example, screening products may identify diseases for which no treatments exist, no effective treatments exist, or for which costs far exceed documented benefits of interventions. They may not be reliable, sensitive, or specific. They may cause patient anxiety, or present an option that is not acceptable to patients. In addition, they may result in added costs to the healthcare system with limited improved patient outcomes.

The difficulty of measuring the efficacy of screening technologies regarding a reduction in mortality is changing. A June 29, 2011, article in the *New England Journal of Medicine* presents the first findings of reduction of lung cancer mortality using a

screening technology.[177] In addition to presenting a reduction in mortality of 20% using low-dose CT, the article also addresses the topic of the appropriate comparator for evaluating a screening technology: an alternative screen or community practice. In the field of HTA, the comparator is a significant issue in order to be certain that the technology meets the HTA criteria that the technology can be used effectively outside of the research setting. "Chest radiography was chosen as the screening method for the control group because radiographic screening was being compared with community care (care that a participant usually receives) in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial."[178]

Today, screening products are a focus of private sector innovation. As a result, in order to encourage the development of screening products with a meaningful health impact to patients, physicians, and health systems, the use of relevant criteria to evaluate these technologies using HTA is important. While HTA findings will likely continue to be diverse given the different focuses of the organizations providing the analysis, use of relevant criteria for HTA of screening technologies will be helpful to manufacturers of these products.

---

[177] Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. The National Lung Screening Trial Research Team. June 29, 2011. http://www.nejm.org/doi/full/10.1056/NEJMoa1102873?query=TOC&#t=article. Accessed July 1, 2011. "In the NLST [National Lung Screening Trial], a 20.0% decrease in mortality from lung cancer was observed in the low-dose CT group as compared with the radiography group. The rate of positive results was higher with low-dose CT screening than with radiographic screening by a factor of more than 3, and low-dose CT screening was associated with a high rate of false positive results; however, the vast majority of false positive results were probably due to the presence of benign intrapulmonary lymph nodes or noncalcified granulomas, as confirmed noninvasively by the stability of the findings on follow-up CT scans."

[178] *Ibid.*

Key Findings

Key findings based upon input from the selected HTA leaders interviewed in this research study in three specific areas provide assistance in shaping more robust criteria for HTA of screening products.

1. Health Technology Assessment at Your Organization

> ***Finding 1: In the experience of the 12 respondents, HTA is incorporated in their organization's mission and vision statements and provides a source of funding for HTA programs.***

All 12 respondents were experts in the field of HTA. The data demonstrated a high expertise rating of 8 or 10 on a scale of 1, not at all familiar with HTA, to 10, extremely familiar. In 6 (50%) respondent answers the organization had HTA as a part of both the mission and vision of the organization.

No respondents indicated that HTA was the sole purpose of their organization. Typically, HTA was a part of the overarching organization. Ten respondents indicated that their organizations had HTA programs since before 2000. All 12 respondents have completed at least 30 HTAs since 2000. Similarly, 5 respondents strongly agreed and 1 agreed that HTA is a source of research funding for the organization. Further, 9 respondents indicated that the HTA department has a specific budget within the umbrella organization.

Thus, the interview data indicates that experienced HTA organizations have integrated HTA into the overarching purpose of the organization. While organizations that offer HTA often have diverse portfolios of services beyond HTA, they have specific budgets to conduct HTAs. In addition, these organizations have funding support for their HTAs that are renewed on an annual basis or are in place for several years at a time.

These professional organizations have both a sustainable internal organizational model and financial incentive to receive future funding in exchange for providing high quality HTAs. As a result, this is an environment that would respond well to implementing a best practice to evaulate different criteria that are specific to screening technologies. This would help to ensure high quality HTAs that will in turn influence recurring funding. Therefore, the development of best practices for HTA of screening technologies would fit within the organizational strategic and funding frameworks of existing HTA organizations.

Finding 2: In the experience of the 12 respondents, HTA is not always linked to coverage decisions and changes to healthcare policy.

All 12 respondents indicated that the HTAs that they conduct serve to "synthesize and summarize information." While some HTA may be used to help a larger health system make a coverage decision, the core focus of the HTA is to present a judgment about a particular product based on existing data and evidence. However, given that 9 out of 12 respondents (75%) indicated that the HTA criteria are different between therapeutic and screening products, to some degree the HTA for screening technologies are more qualitative in nature than perhaps the respondents would like to admit. The end-user of the HTA: payer, government, or other decision-making authority, can use the HTA as the basis for making an informed decision about whether or not to cover a specific technology for a specific patient group.

Several respondents indicated that their HTAs often shape healthcare policy in terms of patient access to a particular technology. In many cases, these assessments affect significant patient populations such as the Medicare population, covered lives of a

significant national private payer, and government health systems for provinces and countries.

Thus, while the strict function of the HTA is to present an unbiased set of facts about a product, a secondary outcome of an HTA may be a coverage decision that has a direct link to patient access to a technology that is reviewed. According to a majority of the respondents, the HTA may, and often does, influence health policy. The literature review indicated that as policies differ, the overarching health policy can become fragmented. However, as these organizations serve different purposes and patient populations, the fragmentation of policy helps patients to be able to choose options that organizations offer that are most relevant to them. As such, the danger of fragmented health policy becomes a positive, rather than a negative, attribute.

Best practices that enable the evaluation of relevant criteria to guide HTAs of screening technologies could help to create better coverage policies for specific patient populations. Consistency in coverage among stakeholders that serve specific patient populations could help to stabilize a payer's patient population. Different payer policies provide patients with choices to participate in plans that provide them with access to specific technologies. This ability to choose may positively influence the amount of time a patient decides to be a member in a payer's plan. As such, a best practice that provides the ability for a HTA organization to use specific criteria geared toward the assessment of screening technologies could enhance coverage for specific patient populations.

> ***Finding 3: In the experience of the 12 respondents, HTA criteria vary depending on the assessment and are available in published reports; however, metrics to measure the actual effectiveness of the HTA do not exist.***

All 12 respondents were asked to share the criteria for HTA at their organization. None provided specific criteria. Seven respondents indicated that their criteria varied because different types of assessments have different requirements. As such, it would be impossible to set forth a specific set of criteria. One respondent commented:

> The political environment wants more flexibility than you would want through, or expect through, a strict HTA. The evidence is but one of the factors that is considered.

Despite the lack of specific criteria, 10 respondents indicated that the criteria and data requirements were published and available in their reports. Seven respondents indicated that feedback from outsiders is accepted. However, 3 did not accept feedback, and two did not provide comment regarding whether or not feedback is accepted. None of the respondents had explicit metrics used by their organization to measure the effectiveness of an HTA.

Thus, while HTA criteria may differ depending on the nature of the assessment, ultimately the criteria are public. However, after the assessment is complete, and possibly affecting coverage decisions and influencing health policy, there are no metrics in place to measure the overall efficacy of the HTA. There was no correlation between influencing coverage decisions and continued funding of HTA organizations.

Criteria specific to HTA of screening technologies would serve to make evidentiary and data requirements transparent to innovators and patients. One outcome of this change might be to influence the variability in the final assessment between different HTA organizations. However, given the different focuses and funding sources of the organizations, variability may continue. Regardless, as demonstrated by the different HTA outcomes for CT colonography, more robust criteria for HTA of screening

technologies could provide a means to consistently evaluate the overall effectiveness of the screening technology in a specific patient population. This would provide access to this technology to a limited subset of patients. Under the current model, because the criteria do not address the needs of screening technologies, the HTA has become a means to deny access to all patients.

> ***Finding 4: In the experience of the 12 respondents, technologies selected for HTA are based on client requests and not necessarily the utility of the product.***

Eight respondents indicated that technologies are selected for assessment based on requests made by clients. These clients fund the organization to conduct the HTA. Examples of HTA clients include: larger government agencies, healthcare systems, managed care organizations, hospital systems, and government agencies. Controversy surrounding a product and whether a product presents a national issue are two of the drivers that can influence requests for HTA.

Best practices that promote the evaluation of criteria specific to HTA of screening products would help to direct client-driven investment in HTA to products that are most helpful to a specific patient population. Development of detailed selection criteria to help determine which screening products might have a broad impact on patient populations, no matter how select or small, would serve to select out screening technologies that do not have a sufficient sensitivity, specificity, or other benefit to merit review. In so doing, this would focus HTA and client resources on screening products that would deliver a more meaningful result to patients and could be used to educate health consumers.

2. Value and Attainability of Outcomes Evidence

Finding 1: In the experience of the 12 respondents, screening products may identify disease and risk factors earlier, but also may increase payer costs with limited patient outcomes.

The data for this particular finding built upon several outcomes from the interviews. First, 8 respondents strongly agreed and 2 respondents agreed that screening products might identify disease and risk factors earlier. Similarly, 7 respondents strongly agreed and 4 agreed that screening products might identify risk factors that never progress to disease. Lastly, 4 respondents strongly agreed and 6 agreed that a screening product might increase payer costs with limited patient outcomes.

While a screening product may have utility in identifying a disease or risk factors, it may also identify risk factors that never progress to disease. Thus, the value of the screening product may be outweighed by the increased costs to a payer or patient for a limited patient outcome. Criteria specific for HTA of screening products would help to focus criteria and data requirements to identify those screening products that have a significant patient outcome. Products with poor outcomes would universally, and consistently, not be recommended by the assessment.

Finding 2: Among the 12 respondents, HTA criteria are sometimes different for therapeutic products and screening products and a hierarchy of evidence exists.

Eight respondents indicated that HTA criteria are different for therapeutic and screening products. Four respondents indicated that the criteria are not different; however, the manner in which the evidence is handled might be different. In addition, 8 respondents indicated that there is a hierarchy of acceptable data. Five of the 8 respondents indicated that randomized controlled trial was at the top of the data hierarchy. One indicated that a hierarchy did not exist.

Thus, depending upon the organization, the HTA for a screening product might require randomized controlled trial data. While at another organization, less rigorous evidence might be permissible. In the case of the more rigorous standard, a screening product may never become available to patients because of the cost involved in conducting a randomized controlled trial, or the time involved to collect the data over time to determine if the disease ever presented in the patient. Effectively, this would eliminate a screening product with positive patient outcomes from being available at all. This finding is consistent with the literature review in that the value and obtainability of outcomes evidence to evaluate screening products in HTA is not clear.

The creation of best practices to evaulate criteria specific to screening products would help clarify evidence criteria and may mitigate evidentiary standards that may not be relevant for these technologies. For example, a randomized controlled trial may not be a reasonable evidence standard for certain screening products. Further, in addition to using evidentiary standards that may not be applicable, answers from the respondents indicate that there is inconsistency in the hierarchy of evidence used by HTA organizations. This inconsistency could prevent helpful screening products from becoming available because the evidence standard is too rigorous, or not possible.

A best practice for selecting criteria relevant to specific screening products would effectively set a transparent standard of evidence based upon specific patient population factors that affect disease such as age, gender, and national origin. This would help to focus the HTA on relevant evidence that is directly related to the affected patient population resulting in a more responsible use of screening products, improved patient access, and consistent health policy regarding screening technologies. Further, this

would address the gap identified in the literature review regarding the importance of considering the patient's experience.

3. Evaluation of Cost Effectiveness of New Screening Technologies

Finding 1: Among the 12 respondents, cost is an assessment measure that should be considered in HTA.

Six respondents strongly agreed and 2 agreed that cost should be added as a measure in HTA. However, several respondents drew a distinction that cost is not an evidence measure it is an assessment measure. Respondents felt that cost should not be a sole consideration when assessing a technology. However, it can be a useful tool in determining the cost versus benefit of a particular product. This finding is consistent with the literature review finding that the value of cost and cost-effectiveness is not clear.

This section of the interview prompted a great deal of additional comment from all respondents. One respondent located in the United States commented regarding the inclusion of cost as an input and the responsibility of the final user of an HTA to contemplate cost-effectiveness:

> Cost is sometimes incorporated in our technology assessments, but not cost-effectiveness. It is up to the final user to do the final step of incorporating cost analysis themselves into the overall evaluation/assessment…we have limits on the types of information we can provide. From a pure speculation standpoint, yes, cost inclusion is inevitable. I think we will need to wait for another Presidential election cycle or other Congressional changes before we will see it happen. And once it does, it will need to be included carefully.

Another respondent commented on the importance of identifying the cost to a specific population and whether cost could be used as a way to disinvest in technology when it is shown not to be effective via HTA:

> I think that one of the areas where cost effectiveness (CE) diverges between the US and other countries, as in Europe, where resources for healthcare are centralized. In this country [the US] you have to think of CE for whom, for which specific population? In other countries, the impact is more on the overall system. Here, the specific populations to consider include Medicare, children and schools, Medicaid…The other question is disinvestment. Should HTAs be used to disinvest previously approved technology because it has been shown not to be effective?

One of the respondents from outside the US commented on the insufficiency of Quality Adjusted Life Years (QALYs) as a measure of cost:

> I am an economist myself, and I strongly believe the economics have to be involved. With economics, comes, by definition a link to outcomes. The usual link is quality of life, specifically cost per QALYs gained. We feel there is a strong need to find something stronger than QALYs but that is not easy. The first thing we need to develop is that something needs to be calibrated, looking very much on age.
>
> With QALYs, let me use an example. So, at the San Francisco General Hospital, there is a patient, who is a murderer and he is in the emergency room. He had just killed school children at a local school, and is 35 years old. At the same time, the ambulance is coming in with a young man who has served with Médecins sans Frontières, one year older than murderer, but suffering from life threatening infectious disease. Who do you treat?
>
> What decision tools do we have? If we use QALYs, then we choose murderer because he is younger. But this goes against society's values, where we would prefer the doctor to the murderer. QALYs are too ice cold. It doesn't discriminate in these types of societal circumstances. QALYs shouldn't be the primary tool for making decisions.

> But it is a straight-forward measure. We should discuss every case, and make age group decisions, which would inform intentions of use.

Regarding the consideration of cost for screening technologies, one respondent commented:

> When you talk about evidentiary standards, it is around risk communication, how you portray risk, and how people interpret those different levels. How distorted willingness to pay gets when you are at low risk. Has to do with perceived value and reaction (favorable or not) to baseline risk and changes in risk.

A component of cost should be considered for HTA of screening technologies. Cost could help to help end-users of HTA for screening technologies understand whether or not they can afford to adopt a particular screening technology for a particular patient population. For example, if a part of the best practice is to further refine the patient population intended for the technology by examining age groups and the burden of the cost of the technology to those age groups, more educated decisions about who would benefit the most from a particular screening product would become more transparent.

In the example of CT colonography, one of the concerns raised in several of the published HTAs was whether or not the discovery of extra-colonic findings would require follow up to determine if the finding was something that needed to be addressed or not. In this case, using the technology in an older population who are more prone to colorectal cancer would help to limit the possibility for overuse of a screening product with limited positive patient outcomes. Understanding what age patient population would benefit the most from the use of screening technology might reduce the cost burden of the technology by using it on those patients that would benefit the most. As a result, a cost measure would be helpful to include for HTA of screening products.

Conclusion

The findings of this research study indicate that the purpose of an HTA is to evaluate existing data and evidence to make a statement about a particular product. HTAs are used as tools by agencies and companies such as: health care plans, managed care organizations and hospitals to decide whether their patient populations should have access to a particular product. As a result an HTA may directly influence payer coverage decisions and, as such, indirectly influence the creation of broader health policy across stakeholders who manage patient populations, as well as affect patient access to care.

Access to healthcare is a changing environment in the United States. In the 1950s, healthcare benefits were focused on the delivery of hospital inpatient care. With the creation of the Medicare program in 1965, new settings of care emerged that required new methods to compensate for care provided in the hospital outpatient and community based settings. Payment mechanisms for these new sites of service have evolved over time as well – transitioning from traditional fee for service models to those based on length of stay or bundled payments. Payments for products used during the course of care have similarly evolved. Itemized payment is almost non-existent as new mechanisms bundle product payment within the total payment for the care of the patient.

These changes to health insurance systems are relevant to the evolution of HTA. In the 1960s, HTA began as a way to distinguish technologies that were beneficial to patients from those that might present and unintended harm if a technology is adopted before understanding whether or not it was safe. Today, HTA is used to drive patient access and payer coverage policy. These mechanisms have a direct effect on the

innovation of new products. If it is unclear or difficult to provide access to specific technologies to patients, innovation in these areas will slow.

Criteria specific to the evaluation of screening products would positively impact HTA stakeholders such as HTA organizations, their clients, and patients, as well as technology innovators. Best practices designed to help HTA organizations choose criteria that are focused on screening technologies will help to identify whether relevant patient populations for the technology exist. In so doing, levels of evidence and data requirements would be more transparent to screening technology innovators and patients. Cost should be a part of the assessment to understand the cost and benefit of using the product in specific patient populations for appropriate clinical decision-making.

As a result, best practices for screening technologies would help provide a relevant channel to evaulate the safety and appropriateness of screening products for specific patient populations via HTA. This is consistent with the evolution of patient care to a more transparent environment in which the patient is more involved in making access decisions.

Limitations of This Research

The components of this research study included the search and assessment of literature, web site content, and primary interviews with HTA experts. The data collected provided a basis to provide answers to the research aims, make conclusions, and develop a plan for change. However, there were limitations to this study that are common to studies that include a review of literature, publicly available web site content, and interviews with technical experts.

A limitation of this study is that mistakes can be made during literature and web site reviews. Key studies and relevant sites can be missed despite best efforts. Links to pertinent parts of web sites can become inactive. These limitations were mitigated by conducting multiple key word searches and by conducting research queries several different ways. These limitations were further mitigated in that the literature that was reviewed was comprehensive and began to cite similar references, indicating that the review was complete.

The search for HTAs of screening technologies was limited by the number of HTAs that were publicly available. This limitation was mitigated by one of the respondents offering to share information they had collected with the interviewer. In addition, respondents shared with the interviewer where information could be found on their websites.

The small sample size for the interviews and the sampling methodology and participation introduced selection bias. This limitation was addressed by involving, when available, more than one person from a particular HTA organization. There are few HTA experts that are extremely familiar with their organization's HTA programs. As a result, the small sample size was anticipated before the study began. The limitation was further addressed by selecting diverse organization types that conduct HTAs. These organizations have a variety of missions, visions, diversity of clients, and funding options. In addition, these organizations were located in very different areas of the United States and address differing patient populations. To manage this limitation, three respondents were included in the study from two different countries to provide additional international perspectives from Canada and the United Kingdom. That said, the limited

sample size is not be representative of the universe of HTAs in all countries. As a result, the findings set forth in this research study may have some applicability to a larger audience; however, some findings may not be generalized to represent all organizations in all locations conducting HTAs.

Another limitation of the research is that the interviews relied heavily on the knowledge and expertise of the respondent. Limitations were introduced by the inability for some of the respondents to be forthcoming with information. Several of the respondents were not able to answer specific questions. For example, a few respondents declined to comment about specific HTA criteria and could not describe specific examples of the differing ways in which data are applied to therapeutic versus screening technologies. Similarly, a few respondents were unable to provide data about the cost of an HTA at their organization due to the varying size and complexity of HTAs conducted. These types of organizationally-specific details are often considered confidential in many cases and therefore cannot be made publicly available. As a result, the absence of information creates gaps of information in the study.

The non-participation of 8 of the 20 respondents invited to participate in the study is another limitation. While this factor did narrow the total number of observations in this study, the 12 respondents did have remarkably similar comments. Due to the geographic distribution of the respondents and the diversity of the types of organizations, it is unlikely that any significant bias was introduced to the study. In addition, a review of publicly available HTAs of CTC colonography for colorectal cancer indicate that the methods employed by non-participating respondents did not reveal any divergence in approach to HTA between participating and non-participating respondents. Thus, the

non-participation of respondents was based more on an individual's time available to participate in the study.

A final limitation is that respondents may unintentionally inject bias into their responses. This bias could come from the manner in which the researcher asked the question. The researcher conducting cognitive tests of the interview guide with three people familiar with HTA prior to the interview process with the respondents mitigated this limitation. This testing helped to refine questions to minimize the possibility of divergence in possible ways to answer the question due to the content of the question being misinterpreted by the respondent. In addition, in one case two people from the same organization were interviewed at the same time. It could be that the differing positions of the respondents in the organization could influence the independence of responses. However, based on the long time, collaborative nature, of these respondents working relationship, it is unlikely that any discussion was stifled.

Any additional study limitations were addressed by comparing the respondents' answers to the documents and websites that they referenced during the interview. When answers provided during the interview were unclear, additional clarifying questions were asked to help narrow the response. Lastly, respondents had the opportunity to review the notes of the interview and make any clarifying edits prior to the analysis of the data. This helped to ensure that the intent of the respondent's comments was clear and well defined.

Chapter 7: Plan for Change: Recommendation for A Standard for HTA of Screening Technologies

This research will form the basis of best practices that can be used by HTA organizations to assess screening technologies. These best practices will help organizations: (1) identify relevant patient populations for specific screening technologies; (2) understand which screening technologies meet their stated organizational goals and objectives; (3) make intentional choices about which screening technologies bring the most value to both patients and payers; (4) direct private sector investment in technologies to improve health outcomes; and (5) provide greater transparency to patients as they select plans that may offer differing levels of access to screening technologies.

Developing best practices is clearly needed, given the findings regarding the diversity in criteria used by key informants respondents regarding CT colonography. Notably, we found differences between standards by which therapeutic and screening products are interpreted, thereby justifying the importance of refined criteria for screening technologies. Presenting these refined criteria in the form of a "best practice" would provide HTA organizations with the flexibility to decide which criteria best serve their mission and responsibility to patients. Specifically, based upon literature review, review of HTA, and interviews with 12 leaders in HTA, five best practices were identified (Table 32).

Table 32. Best Practices for HTA of Screening Technologies

Finding	Thematic Finding
1	Determine the types of data and evidence that are sufficient to support screening reliability, sensitivity and specificity
2	Evaluate data to help identify the appropriate patient population. Take into consideration demographic, clinical and genetic characteristics.
3	Reference the natural course of disease.
4	Ethical considerations such as: the link between the test and actual expression of disease, availability of treatment options if disease is identified, and taking into consideration patient acceptability.
5	Consider the cost.

A discussion of each of these thematic findings follows.

1. ***Determine the types of data and evidence that are sufficient to support screening reliability, sensitivity and specificity.***

One controversy surrounding screening technologies is their ability to be sufficiently reliable, sensitive, and specific. Without these endpoints, the technology could result in elevated costs to the healthcare system and unnecessary risks to patients when seeking to diagnose a disease that is not present (false positive). Several respondents commented about the potential consequence of screening healthy patients, including exposing patients to unnecessary anxiety or treatment. However, reliability, sensitivity and specificity must be balanced against an individual's right to choose whether to be screened for a particular disease. One respondent commented:

> This can be highly contextual. For screenings, you can be offering them, but since people know there isn't any treatment, then it would be their option to use it. If a screening product is highly accurate, highly effective, then patient and clinician preference plays a big role. That outweighs whether the organization says they should have that test or not. It really becomes an individual decision.

Thus, an HTA assessment best practice for screening products should balance detection of disease, treatment of disease, and better patient outcomes. While randomized controlled trial data may be too high of a bar for screening technologies, a

best practice should be able to validate the possibilities of applying differential evidentiary standards to the severity of particular diseases. For example if a disease is particularly severe, the ability to evaluate whether or not a patient might have the early stages of the disease may present more effective treatment options than if the patient waited for the disease to formally manifest itself. Based on this best practice, a patient should have the right to choose whether or not they would like to have the test. As such, the individual and their physician may make informed choices about treatment relative to the reliability, sensitivity and specificity of the test.

2. ***Evaluate data to help identify the appropriate patient population. Take into consideration demographic, clinical and genetic characteristics.***

A best practice for the HTA of a screening technology should consider the characteristics of individual patients (e.g., demographic, clinical and genetic characteristics) so that the product is used in the most appropriate patient population. Doing so will enhance the test characteristics (sensitivity and specificity). Without the ability to segment patient populations, we are at risk for screening patients for disease that does not take into consideration the overall relevance of the test to the particular patient. One respondent suggested that:

> An example of a harmful screening program – test for prostate cancer using PSA [for men] over [age] 64. Many false positives. As a result, men go unnecessarily to surgery and there are many side effects and complications.

Narrowing screening tests to relevant patient populations could make a substantial difference in the manner in which we permit testing. One respondent discussed mammography testing in different age groups:

> Whether you are looking at potential benefit over lifetime, or adverse events of surgery, it becomes a different question

> if evaluating in patients who are pre-menopausal or post-menopausal. If the patient is 70 years old, and there is a concern about cancer, the BRCA mutation, in this case, isn't the underlying cause. To subject a 70 year old with lower tolerance for surgery and multiple comorbidities to invasive surgery, it becomes very different if you're thinking about a 25 year old, who would be a prime candidate for surgery. The bigger risk in that age group would be fertility for example, because of the bilateral oophorectomy.

As a result, a best practice for a screening HTA should be to identify the most relevant patient population for the test. This would help to avoid unnecessary testing, inappropriate testing, and focus the effort on testing in the most relevant patient populations.

3. Reference the natural course of disease.

While a screening product might establish that disease is present, it is important to understand the overall course of the disease. For example, some therapies are not relevant to specific patient within the context of their particular situation. One respondent commented:

> A lot what we screen for, risk factors, biomarkers, we need sufficient evidence to connect to them to the disease condition that you're testing for. Then we need the information to link that it can be influential on decisions and that there is something that the physician can do.

Patient treatment and survival are not always appropriate endpoints for specific patient populations. For example, screening tests could be provided to patient populations that are not relevant and unnecessary patient anxiety can ensue at the expense of a sole focus on treatment and survival. One respondent commented that a negative benefit may come from screening technologies:

> [In a study for mammography w]e followed up with women who [were] positive one way or another – that

> through the screening, they find something that needs to be examined further…we took a cohort of 300 women, and followed over the years, to see when they were free of initial suspicion of "seeing something". We knew the results, and we took the cohort of false positives, to look backwards and see. During this period of time, many of these women fell into depression and had anxiety. We argued that this should be considered a negative benefit of mammography screening, but has not been considered a concern. The only concern is the aggressive treatment and survival.

As innovators develop screening tests for future disease, we will need to consider the natural course of the disease. In so doing, a best practice for HTA of screening products would help identify where the patient is in the natural history of the disease, and conversely, when treatment is not helpful given the overall natural course of the disease. By understanding the broader context of the natural development of a particular disease, HTA for screening tests could be refined to create statements of fact regarding when treatment is most effective or not warranted.

4. ***Ethical considerations, such as: the link between the test and actual expression of disease, availability of treatment options if disease is identified, and taking into consideration patient acceptability.***

A best practice for the HTA of a screening product should consider data that helps to establish a link between the screening product and the actual expression of disease. One of the key considerations expressed by respondents was that if no treatment for a screened disease exists, then the product should not be used.

> Some proponents feel that information for information's sake can be good. But it can do harm if you don't understand meaning of information and implications on disease. If there isn't any action that flows from it, then can do more harm than benefit. It creates anxiety. The downstream of positive sequela, is retesting. Negative test results may give false reassurance. This is a hard concept to explain to public. We haven't done a good enough job of

> explaining of what makes a good screening test valuable and why.

The level of analytic precision of the screening product presents a higher standard. Specifically, the link between the test and the formal expression of disease is particularly important. One respondent commented:

> Unlike treatment, where questions are more simplistic, [for example], does it work, how well does it work, is it safe, how may it compare with an alternative approach…If you are talking about molecular diagnostic testing, you are adding another dimension, including things that might influence expression of gene of allele that you're testing for…What is the link between what you are testing for and actual expression of disease?

Patient acceptability of the elements necessary to conduct the screening test are also important. One respondent commented:

> Everyone agrees that [colorectal cancer testing] is a very important technology. Part of the process includes submitting a fecal sample, which the patient need[ed] to mail for analysis. There are no ethical issues presented, but in this case, not everyone wants to process and mail their feces.

If the steps of a test, regardless how simplistic, are not easily accepted by the patient, then the overall value of the test is inconsequential because there is a barrier to use. In this case, the test was less invasive and the patient could ensure privacy; however, the cultural context was not acceptable and the patient compliance with the test was not strong.

As a result, a best practice for HTA is for the assessment to take into consideration that testing for the sake of testing is not a beneficial concept. The best practice should not be solely focused on treatment and survival. Rather, a part of the assessment should be centered on the identification of the most appropriate patient

population. This would help to avoid a potential for being the source of patient anxiety. Lastly, the best practice should take into account the patient's ability to accept the actual process of the test. Without patient willingness, the test is not valuable.

5. ***Consider the Cost.***

Distinguish valuable screening technologies from those that offer added costs with limited improved patient outcomes. Cost should be considered in conjunction with potential benefits. One respondent commented:

> There is the whole issue of false positives and false negatives to consider. You also need to look carefully at the cost, of course, related to the potential benefits.

A majority of respondents indicated that cost should be an assessment measure in HTA. An evaluation of cost helps health systems plan their investments. In addition, transparency of costs helps patients choose whether or not they want a particular screening test.

A best practice for HTA of screening technologies should include component of cost in order to inform both of these stakeholders. This component could be in the form of a cost-effectiveness study to understand the cost of the screening product within the broader context of the treatment of the disease. This view into cost could help to identify valuable screening products with positive patient outcomes from those products with high cost and little benefit.

Implementation of HTA Best Practices for Screening Products

Issues of coverage, access, and affordability are important components of the healthcare policy environment in 2011. Coverage decisions affect whether or not a patient will have access to a particular product, which in turn impacts overall

affordability to both the patient and payer. The respondents in this study indicated that the organization conducting the HTA might not be directly responsible for making a coverage decision in all cases. However, HTAs are often used by health plans and hospital groups as a means to make coverage decisions for their members. As a result, once a screening product receives a positive HTA and is deemed eligible for coverage by the payer, follow up questions exist relative to appropriate patient access and affordability. By implementing the best practices identified in this research study, HTAs play a direct role in helping to establish a rational way for screening products to be covered by payers so that appropriate patient populations can have access to them. If implemented correctly, screening products will help payers to manage healthcare expenses by identifying patients for whom a screening test might have the most value.

How should the HTA best practices for screening products be effectively implemented by payers in today's healthcare environment? The respondents in this study indicated that there is a concern regarding appropriate access to screening products when a payer's patient population has unrestricted access to all covered products. Two examples from this study included the overall value of both PSA tests and mammograms to patients and payers when used in elderly patient populations. The respondents stated that while these tests have a high value in younger populations by identifying disease early, their utility typically wanes significantly when used in elderly populations for whom treatment options are limited due to patient age and ability to tolerate treatments. Several respondents cited concerns related to increased patient anxiety and cost to the payer as follow up tests and procedures are ordered in response to the results of the PSA or mammogram. These respondents felt that there is a definitive link between the value

of the test to both the patient and the payer. As these two examples demonstrate, the value of a screening product is not fully realized when a coverage policy assumes that all patients are the same and therefore should have universal access.

Value-based insurance designs are an ideal mechanism for payers to realize the benefit of HTA best practices for screening products. The principle of value-based insurance design is the value of a technology (medication, program, screening test) varies within and across patient populations.[179] As a result, access to screening products should be based upon when the clinical benefit exceeds the cost in specific patient populations. Access should not be driven by the uniformity of patient out of pocket expenses.[180] Value-based insurance design provides the mechanism for payers to incentivize services that are most valuable to the patient. This helps to direct access to those products that demonstrate value while considering cost. Similarly, it motivates providers to either decrease cost or increase quality to justify a higher price.[181] Thus, "[e]fficiency would promote the use of 'valuable' interventions whose expected net clinical benefits justify the associated expenditure and limit access to those services whose costs exceed the expected clinical gain. This is the fundamental paradigm of cost-effectiveness analysis."[182]

Payers with value-based insurance designs could use several strategies to manage the use of screening products supported by HTA to make decisions about coverage,

---

[179] Chernew, M.E., Rosen, A.B., Fendrick, M. (2007). Value-based insurance design. *Health Affairs*; 26:2, w195-w203.

[180] *Ibid* at w195.

[181] Ginsburg, P.B. (2007). Shopping for price in medical care. *Health Affairs*; 26:2, w208-w216.

[182] Chernew, M.E., Rosen, A.B., Fendrick, M. (2007). Value-based insurance design. *Health Affairs*; 26:2, w195-w203.

access and affordability. As a starting point, payers should get involved in the development of screening products by sharing with screening product manufacturers what types of products would be most valuable to specific patient populations in their membership. Payers are in the best position to know what types of screening products would be of the most use to their members. As a result, by proactively interacting with the inventors and manufacturers of screening products, the innovation will match the need.

Once the need is established, the HTA for a screening product would employ the best practices to identify appropriate patients while considering cost. If coverage for the screening product is established, value-based insurance designs would be the mechanism for directing screening products to the patient populations for whom the products will have the most value. Patient demographics in the coverage policy would help physicians to identify which patients should be proactively screened. Alternatively, a patient that is not in the covered population could choose to pay the out of pocket costs associated with access to the particular screening product. In this way, appropriate patients are covered and have access, while those patients who choose to have the test but are not a part of the covered patient population could have access via a higher out of pocket, thereby limiting financial exposure to the payer.

For screening products with less certain value propositions absent a lengthy observation period, payers could employ different variants of pay for performance and risk-share models. For example, in a strict pay for performance model, screening products would only be covered each time the test reliably establishes that a patient has a disease. Alternatively, in a risk-share model a screening product would be partially

covered by both the manufacturer and the payer while clinical evidence and cost effectiveness data are gathered to support whether or not the product has a value to specific patient populations within the payer's membership. Currently, Medicare uses coverage with evidence development[183] for newer technologies that show a promise of value. This shared risk model would help to stimulate innovation by mitigating the cost to the manufacturer associated with a lengthy study and data collection, while providing payers with access to new screening products for their members. Lastly, for those screening products with limited data and/or unclear value, patients could opt to pay entirely out of pocket for access.[184]

The practical implementation of the best practices for HTA of screening products can benefit payers that employ value-based insurance designs. These types of designs help payers to identify screening products in which the clinical benefit exceeds the cost. As screening products are adopted, they may establish varying degrees of efficacy in different patient populations. HTA for screening products could become a more active tool to establish coverage which would be managed through value-based insurance design and cost-effectiveness analysis to provide patients access while managing affordability concerns. Additional studies will need to be conducted relative to the feasibility and cost associated with identifying relevant patient populations within payer memberships and establishing whether the clinical benefit of a particular screening product exceeds the cost.

---

[183] Centers for Medicare and Medicaid Services. Coverage with evidence development. https://www.cms.gov/CoverageGenInfo/03_CED.asp Accessed September 2, 2011.

[184] Brennan, T., Reisman, L. (2007). Value-based insurance design and the next generation of consumer-driven health care. *Health Affairs*; 26:2, w204-207.

Appendix 1: Recruitment Document

Study #: 11-0624

As you may know, screening technologies may be able to help identify whether or not a patient has a specific disease in order to inform treatment decisions before symptoms occur or verify the cause of symptoms. It is believed that finding a disease early might improve a patient's prognosis. However, there is not consensus as to whether the discovery of a disease in an early phase may ultimately result in better patient outcomes. Regardless, a significant amount of private sector innovation is occurring to develop new technologies that can identify a disease early. As such, it is unclear how far down the clinical pathway one must go in order to validate that a screening product actually helped the patient. The purpose of this research study is to identify whether or not agencies that conduct health technology assessments in the United States are employing similar standards of evidence to evaluate products that screen for disease. The results of this research study will be used to develop a best practice for conducting these types of health technology assessments.

Your organization has been identified as a leader in the area of health technology assessments. In addition, you have been identified as someone with expertise in this area. While there is a great deal of speculation about the levels of evidence required to support a HTA of a screening product, there does not appear to be a common standard. Information from this research study will be used to try to create such a standard.

During the interview, you may choose not to participate, decline answering any question, or stop at any time. However, your response, if you do participate, will be valuable to the results of this research study. The information collected in this research study will be kept confidential. Your specific answers will not be attributed to you or your organization. The interview results will only be used in summary form to discuss the suggested levels of evidence for these technologies and to form specific recommendations for organizations wishing to create a best practice for evaluating these new technologies.

I am the Principal Investigator for this research study. My faculty advisor is John Paul, PhD. Dr. Paul is a Clinical Associate Professor at University of North Carolina, Chapel Hill, Gillings School of Global Public Health. His telephone number is: 919-966-7373 and his email is: paulj@email.unc.edu

Appendix 2: Interview Guide

Interview Guide

Health Technology Assessment Evidence Criteria: What Types of Evidence Should be Presented for Products Used to Screen for Disease in the United States?

Interviewee Name: _____

Interviewee Title: _____

Organization Name: _____

Date of Interview: _____

Interview Start: __:__

Interview End: __:__

Duration of Interview: _____minutes

Introduction

Good morning/afternoon. My name is Nancy McGee, and I am a student in the Executive Doctoral Program in Health Leadership at the University of North Carolina at Chapel Hill Gillings School of Global Public Health.

Thank you for your interest in participating in this research study entitled: Health Technology Assessment Evidence Criteria: What Types of Evidence Should be Presented for Products Used to Screen for Disease in the United States. I am the Principal Investigator for this research study. My telephone number is: 415-279-4448 and my email address is: nancymmcgee@gmail.com. My Faculty Advisor is John Paul, PhD he is a Clinical Associate Professor at the University of North Carolina, Chapel Hill, Gillings School of Global Public Health. Dr. Paul's telephone number is: 919-966-7373, and his email address is: paulj@email.unc.edu All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject you may contact, anonymously if you wish, the Institutional Review Board at 919-966-3113 or by email to IRB_subjects@unc.edu.

As you may know, screening technologies may be able to help identify whether or not a patient has a specific disease in order to inform treatment decisions before symptoms occur or verify the cause of symptoms. It is believed that finding a disease early might improve a patient's prognosis. However, there is not consensus as to whether the discovery of a disease in an early phase may ultimately result in better patient outcomes. Regardless, a significant amount of private sector innovation is occurring to develop new technologies that can identify a disease early. As such, it is unclear how far down the clinical pathway one must go in order to validate that a screening product actually helped the patient. The purpose of this research study is to identify whether or not agencies that conduct health technology assessments in the United States are employing similar standards of evidence to evaluate products that screen for disease. The results of this research study will be used to develop a best practice for conducting these types of health technology assessments.

Your organization has been identified as a leader in the area, of health technology assessments. Your participation in this interview is key and very much appreciated. While there is a great deal of speculation about the levels of evidence required to support a HTA of a screening product, there does not appear to be a common standard. This research will be used to try to create such a standard.

You may choose not to participate, decline answering any question, or stop at any time. However, your response, if you do participate, will be valuable to the results of this research study. The information collected in this research study will be kept confidential. Your specific answers will not be attributed to you or your organization. The interview results will only be used in summary form to discuss the suggested levels of evidence for these technologies and to form specific recommendations for organizations wishing to create a best practice for evaluating these new technologies.

'egin interview of administrator or manager of a health technology assessment rganization with Section A. If interviewing a panelist or researcher, skip ahead to ections B and C.

[If interviewing an administrator or manager:]

As an administrator or manager of a health technology organization, I will ask you questions about:

A. Health technology assessment at your organization

[If interviewing a panelist or researcher:]

As a health technology assessment panelist or researcher, I will ask you questions about:

B. Value and attainability of outcomes evidence
C. Evaluation of cost-effectiveness of new screening technologies

I anticipate that this interview will take one hour. The interview is not being taped. I will be taking handwritten notes as we talk. At your request, you may review the summary of this interview, as well as the relevant portion of the dissertation that contains the interview content, in order to assure sufficient accuracy, attribution, and confidentiality.

Do you have any questions? May we begin?

A. Health technology assessment at your organization

As an administrator or manager of a health technology assessment organization, I would like to ask you several general questions about health technology assessment at your organization.

1. On a scale of one to ten, how familiar are you with the HTA program at your organization?

Not At All Familiar Extremely Familiar
1 2 3 4 5 6 7 8 9 10

[If response is 4 or lower, thank the interviewee for their participation and ask if they can recommend another individual in their organization that might be more familiar.]

2. There are a number of **critical components for the development of health technology assessment programs** at organizations such as yours. The purpose of this question is to understand the most important components for development of HTA programs at your organization. We would like to know how would you rank these critical components. In this section of the interview, I will read a statement to you, please respond whether you agree or disagree with the statement. We will use a 5 point Likert Scale in which 1 is "strongly disagree", 2 is "disagree", 3 is "neither agree nor disagree", 4 is "agree", and 5 is "strongly agree". Are you ready? Let's begin.

Critical components for the development of HTA programs	**5 Point Likert Scale**				
1. HTA is a part of the mission statement of the organization	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
2. HTA is a part of the vision for the organization	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
3. HTA is a source of research funding for the organization	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
4. The HTA department has a specific financial budget within the organization	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
5. HTA is the sole purpose of the organization	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
6. HTA drives coverage decisions at the organization	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
7. Other:					

3. Are there stated goals or objectives for the HTA program(s)? If so, what are they?

4. When (month and year) did you implement your HTA program?

5. How many HTAs has your company/agency conducted since 2000?

   <10 11-20 21-30 30-40 >40

6. What are the criteria for HTA at your organization?

7. Are the criteria and data requirements published or contained/included in any internal reports that are available?

   a. Is it possible to obtain copies of the criteria?

8. How does your organization select technologies to assess?

9. Can you describe an HTA decision your agency made?

a. Once the assessment was complete, did policies change over time?

b. Did feedback from your providers or patients or manufacturers cause you to change or re-evaulate the decision?

c. Do you use a committee structure to evaluate information? If yes, how is the committee formed?

d. Once a decision is made using HTA, what is the process of disseminating decisions? Are decisions linked to reimbursement or are decisions guidelines?

10. Can you describe an HTA decision you changed because new information came to light?

    a. If so, what was the new information that caused a change and how did you find it?

11. What metrics have you established to measure the effectiveness of HTAs?

12. Do you have ongoing or multi-year budget commitment for HTA assessments? For how many years?

13. Do you know what the average cost of conducting an HTA is for your organization?

14. Is there anything else that you think other organizations tackling this issue should know in evaluating screening products, either before they begin, during the analysis, or when the analysis is complete?

Thank you for your time. Your input and insights will be invaluable to this research study. You will be provided with a summary of the final document, expected in October 2011.

organizations. Health technology assessment <u>administrators and managers</u> will not answer this section.

B. Value and attainability of outcomes evidence

As a health technology assessment panelist or researcher for a health technology assessment organization, I would like to ask you questions about the value and the ability to attain outcomes evidence when evaluating screening technologies.

1. Screening technologies may be able to help identify whether or not a patient has a specific disease in order to inform treatment decisions before symptoms occur. **Controversies** about screening technologies exist. In this section of the interview, I will read a statement to you, please respond whether you agree or disagree with the statement. We will use a 5 point Likert Scale in which 1 is "strongly disagree", 2 is "disagree", 3 is "neither agree nor disagree", 4 is "agree", and 5 is "strongly agree". Are you ready? Let's begin.

Screening product controversy	5 Point Likert Scale				
1. May be able to <u>identify</u> disease earlier	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
2. May identify risk factors that <u>never progress</u> to disease	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
3. May be able to <u>decrease payer costs</u> with positive patient outcomes	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
4. May <u>increase payer costs</u> with limited patient outcomes	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
5. May identify diseases for which there is <u>no treatment</u> available	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
6. Other					

2. Are the **HTA criteria** different between therapeutic products and screening products?

   a. If so, how?

3. What kinds of data are acceptable in order to support a particular HTA criterion?

4. Is there a hierarchy of acceptable data (e.g.: randomized clinical trial data v. adaptive trial data)?

    a. If so, what is the hierarchy of data?

5. Are the **data requirements** different between therapeutic products and screening products?

    a. If so, how?

6. Please rank your level of agreement or disagreement with the following statements:

    a. My organization tends to avoid assessing technologies that screen for disease because it is difficult to agree regarding the appropriate length of the observation period.

1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree

    b. My organization tends to avoid assessing technologies that screen for disease because the cost of these technologies is not sufficiently understood.

1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree

c. My organization tends to avoid assessing technologies that screen for disease because it is difficult to select the appropriate patient population.

1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree

C. Evaluation of cost-effectiveness of new screening technologies

As a health technology assessment panelist or researcher for a health technology assessment organization, I would like to ask you questions about the evaluation of cost-effectiveness of new screening technologies.

1. In Europe, HTA conducted by government payers employ **cost as an evidence measure**;[185] while in the US, cost has typically not been included in assessments conducted by public payers. In this section of the interview, I will read a statement to you, please respond whether you agree or disagree with the statement. We will use a 5 point Likert Scale in which 1 is "strongly disagree", 2 is "disagree", 3 is "neither agree nor disagree", 4 is "agree", and 5 is "strongly agree". Are you ready? Let's begin.

Cost as an evidence measure	**5 Point Likert Scale**				
1. Cost should be added as an evidence measure for HTA	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
2. In general, cost is an unstated evidence measure for all HTAs	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
3. Cost is an unstated evidence measure for HTA at my organization	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
4. Cost is not a relevant evidence measure for HTA	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree
5. Other	1 Strongly Disagree	2 Disagree	3 Neither Agree nor Disagree	4 Agree	5 Strongly Agree

Closing

Thank you again for your time. Your input and insights will be invaluable to this research study. You will be provided with a summary of the final document, expected in October 2011.

[185] NICE: Our Guidance Sets the Standard for Good Healthcare; NHS National Institute for Health and Clinical Evidence. Page 9. 30 December 2008. http://www.nice.org.uk/aboutnice/?domedia=1&mid=EE5AA72F-19B9-E0B5-D4215C860E77FD2E Accessed February 22, 2009.

References

Banta HD, Luce BR. (1993). *Health Care Technology and Its Assessment: An International Perspective*. New York, NY: Oxford University Press.

Birch, S., & Ismail, A. I. (2002). Patient preferences and the measurement of utilities in the evaluation of dental technologies. *Journal of Dental Research, 81*(7), 446-450.

Bodenheimer, T., & Fernandez, A. (2005). High and rising health care costs. part 4: Can costs be controlled while preserving quality? *Annals of Internal Medicine, 143*(1), 26-31.

Brawley, O. W., & Kramer, B. S. (2005). Cancer screening in theory and in practice. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 23*(2), 293-300.

Brennan, T., Reisman, L. (2007). Value-based insurance design and the next generation of consumer-driven health care. *Health Affairs*; 26:2, w204-207.

Bruns, D. E. (2001). Laboratory-related outcomes in healthcare. *Clinical Chemistry, 47*(8), 1547-1552.

Burch JA, Soares-Weiser K, St John DJ, et al. (2007). Diagnostic accuracy of faecal occult blood tests used in screening for colorectal cancer: A systematic review. *Journal of Medical Screening*; 14(3):132–137.

Cook DJ, et al. (1992). Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*. 4(suppl):305S-311S.

Chernew, M.E., Rosen, A.B., Fendrick, M. (2007). Value-based insurance design. *Health Affairs*; 26:2, w195-w203.

Feldman M. (2004). Computed tomographic colonography (virtual colonoscopy) for screening of colorectal cancer. *California Technology Assessment Forum*. June 9, 2004.

Gatto NM, Frucht H, Sundararajan V, et al. (2003). Risk of perforation after colonoscopy and sigmoidoscopy: A population-based study. *Journal of the National Cancer Institute*; 95(3):230–236.

Ginsburg, P.B. (2007). Shopping for price in medical care. *Health Affairs*; 26:2, w208-w216.

Hartmann, K. E., Nanda, K., Hall, S., & Myers, E. (2001). Technologic advances for evaluation of cervical cytology: Is newer better? *Obstetrical & Gynecological Survey, 56*(12), 765-774.

Heresbach D, Chauvin P, Grolier J, et al. (2010). Cost-effectiveness of colorectal cancer screening with computed tomography colonography or fecal blood tests. *European Journal of Gastroenterology & Hepatology*, 22:11, 1372-1379.

Inadomi, J. M. (2008). Taishotoyama symposium barriers to colorectal cancer screening: Economics, capacity and adherence. *Journal of Gastroenterology and Hepatology, 23 Suppl 2*, S198-204.

Johnson CD, Chen MH, Toledano AY, et al. (2008). Accuracy of CT colonography for detection of large adenomas and cancers. *N Engl J Med*, 359(12):1207-17, page 1208.

Liedenbaum MH, Vries AH, Rijn AF, et al. (2010). CT colonography with limited bowel preparation for the detection of colorectal neoplasia in an FOBT positive screening population. *Abdominal Imaging*, 35:6, 661-668.

Pickhardt PJ, Choi JR, Hwang I, et al. (2003). Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *New England Journal of Medicine*; 349(23):2191–2200.

Rex DK, ACG Board of Trustees. (2004). American College of Gastroenterology action plan for colorectal cancer prevention. *American Journal of Gastroenterology*; 99(4): 574–577.

Rogowski, W. (2007). Current impact of gene technology on healthcare. A map of economic assessments. *Health Policy (Amsterdam, Netherlands), 80*(2), 340-357.

Sacher-Huvelin S, Coron E, Gaudric M, et al. (2010). Colon capsule endoscopy vs. colonoscopy in patients at average or increased risk of colorectal cancer. *Alimentary Pharmacology & Therapeutics*, 32:9, 1145-1153.

Schneewiss, S. (2007). Developments in post-marketing comparative effectiveness research. *Clinical Pharmacology & Therapeutics*, 82(2), 143-156.

Schneider, A., Rosemann, T., Wensing, M., & Szecsenyi, J. (2005). Physicians perceived usefulness of high-cost diagnostic imaging studies: Results of a referral study in a german medical quality network. *BMC Family Practice, 6*(1), 22.

Sunshine, J. H., & Applegate, K. E. (2004). Technology assessment for radiologists. *Radiology, 230*(2), 309-314.

Summerton S, Little E, Cappell MS. (2008) CT colonography: Current status and future promise. *Gastroenterology Clinics of North America*; 37(1):161–189.

Super, N. (2004). Medicare's chronic care improvement pilot program: What is its potential? *Issue Brief (George Washington University.National Health Policy Forum), (797)*(797), 1-20.

Tatsioni, A., Zarin, D. A., Aronson, N., Samson, D. J., Flamm, C. R., Schmid, C., et al. (2005). Challenges in systematic reviews of diagnostic technologies. *Annals of Internal Medicine, 142*(12 Pt 2), 1048-1055.

Vermeulen, V., Coppens, K., & Kesteloot, K. (2001). Impact of health technology assessment on preventive screening in Belgium: Case studies of mammography in breast cancer, PSA screening in prostate cancer, and ultrasound in normal pregnancy. *International Journal of Technology Assessment in Health Care, 17*(3), 316-328.

Walsh J. Computed tomographic colonography (virtual colonoscopy) for colorectal cancer screening in average risk individuals. (2009). *California Technology Assessment Forum*. March 11, 2009.

American Cancer Society. Can colorectal cancer be found early? http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-detection Accessed: December 27, 2010.

American Recovery and Reinvestment Act of 2009 (February 17, 2009). http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_bills&docid=f:h1enr.pdf. Accessed: January 17, 2010.

Centers for Disease Control and Prevention. Colorectal Cancer Screening Guidelines. http://www.cdc.gov/cancer/colorectal/basic_info/screening/guidelines.htm Accessed: December 27, 2010.

Centers for Medicare and Medicaid Services. Coverage with evidence development. https://www.cms.gov/CoverageGenInfo/03_CED.asp Accessed September 2, 2011.

CT Colonography ("Virtual Colonoscopy") for Colon Cancer Screening. (2009). Blue Cross Blue Shield Association, Technology Evaluation Center. Assessment Program, Volume 24, No. 1, page 1.

Computed tomographic colonography (virtual colonoscopy). (2005). National Institute for Health and Clinical Excellence. http://www.nice.org.uk/guidance/IPG129 Accessed: November 21, 2010.

Initial National Priorities for Comparative Effectiveness Research, Committee on Comparative Effectiveness Research, Prioritization Board on Health Care Services, page S-9. Institute of Medicine of the National Academies. http://www.nap.edu/catalog.12648.html. Accessed December 6, 2009.

National Cancer Institute, U.S. National Institutes of Health. Colorectal Cancer Screening. http://www.cancer.gov/cancertopics/factsheet/Detection/colorectal-screening Accessed: November 13, 2010.

National Cancer Institute, U.S. National Institutes of Health. Colon and Rectal Cancer. http://www.cancer.gov/cancertopics/types/colon-and-rectal Accessed: November 13, 2010.

Madison Clinic, Harborview Medical Clinic, University of Washington Medical Center. Screening for Colorectal Cancer. http://depts.washington.edu/madclin/providers/guidelines/colorectal.html Accessed: December 27, 2010.

PDQ® Cancer Information Summary. National Cancer Institute; Bethesda, Maryland. *Colorectal Cancer Screening—Health Professional*. Date last modified: 08/26/2008. http://www.cancer.gov/cancertopics/pdq/screening/colorectal/healthprofessional. Accessed: October 3, 2008.

Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. The National Lung Screening Trial Research Team. June 29, 2011. http://www.nejm.org/doi/full/10.1056/NEJMoa1102873?query=TOC&#t=article. Accessed July 1, 2011.

U.S. Preventative Services Task Force. (2008). Screening for colorectal cancer: U.S. Preventative Services Task Force Recommendation. *Ann Intern Med*, 149:627-37, at 631.

Washington State Health Care Authority. (2008). Health Technology Clinical Committee. Findings and Coverage Decision: Computed Tomographic Colonography. Meeting date: February 15, 2008. Final adoption: August 15, 2008.