



Molecular Pathology Testing for the Estimation of Prognosis for Common Cancers

Sreelatha Meleth MA, MA, MS, PhD

Daniel E. Jonas, MD, MPH

RTI-UNC EPC

March 24th , 2015

Co Authors & Funders

- Katherine Reeder-Hayes, M.D.,
- Mahima Ashok, Ph.D., M.S.,
- Robert Clark, Ph.D.,
- William Funkhouser, M.D., Ph.D.,
- Roberta Wines, M.P.H.,
- Christine Hill, M.P.A.,
- Ellen Shanahan, M.A.,
- Emily McClure, M.S.P.H., M.A.,
- Katrina Burson, M.N., B.S.N.,
C.C.R.P.,
- Manny Coker-Schwimmer, M.P.H.,
- Nikhil Garge, M.S.,

Based on research conducted by the RTI-UNC EPC under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD HHS2902007100561.

The findings and conclusions in this document are those of the authors who are responsible for its contents. The findings and conclusions do not necessarily represent the views of AHRQ.

None of the investigators has any affiliations or financial involvement related to the material presented in this report.

Objectives

- To conduct a systematic review assessing 11 molecular pathology tests that might inform estimation of prognosis.
- Our overarching question was whether there is direct evidence that the addition of the results of these molecular pathology tests to traditional prognostic factors changed physician decisionmaking and improved clinical outcomes for adult patients.

Cancer Site	Molecular Pathology Tests Reviewed				
Breast	Mammaprint	Oncotype Dx			
CRC	BRAF	KRAS	MSI	MLH1	Oncotype Dx
Lung	EGFR	KRAS	ALK translocation		
Bladder	Urovysion				

Background – Cancer Incidence and Mortality

- It is estimated that there will be approximately 1.67 million new cases of cancer in 2014
- Cancer is the second leading cause of death in the United States
- Death rates for all cancers are declining. In the last five years overall death rates decreased 1.8% a year for men and 1.4% a year for women. Declines largely due to declines in death rates in the 4 major cancers
- In the last 5 years, lung cancer death rates are down 34% for men, 9% for women; also 34%, 45% and 46% decreased mortality in breast , prostate colorectal cancer respectively.

Impact of Molecular pathology on Ca Mortality

- Advances in molecular pathology have resulted in better understanding of cancer subtypes and development of treatments based on these sub-types.
- E.g. Identification of human epidermal growth factor 2 (HER2) receptor resulted in targeted therapies for breast cancer.
- Advances in molecular pathology have also helped identify tumor characteristics that help predict the prognosis for a patient in addition to traditional markers such as stage and differentiation.

Clarification

MedCac questions 2 & 3 ask about the anti-cancer treatment therapy being guided by these tests. Important to note that these genetic tests are used in two different contexts.

- In one, the tests are used in a specific context of a test/therapy combination, where the test is being used to predict response to a very specific treatment.
- In the second context, the tests are used to estimate the patient's prognosis, and physicians use this prognostic information in a variety of ways (including informing choices from a variety of different treatment options).

CMS requested this report to evaluate the second context. Therefore, studies that evaluate specific test/therapy combinations were excluded from this review

Methods - Overview

- Refined Key Questions (KQs).
- Developed Analytic Framework for KQs.
- Searched databases.
- Systematic Review of the Published Evidence.
- Summarized evidence qualitatively & quantitatively, with a meta-analysis when appropriate.

Methods: EGAPP

- The methods used in this review were based on the recommendations of the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group (EWG).
- The EWG was established in 2005 to develop a systematic process for evidence-based assessment that is specifically focused on genetic tests and other applications of genomic technology.
- The methods developed and recommended by the EWG share elements with many existing processes, such as the USPSTF and the AHRQ Evidence-based Practice Center Program.
- It also recognizes that the gold standard for direct evidence – randomized clinical trials(RCTs) may not be available in the evidence base for these new tests and outlines a process for building a chain of evidence

Chain of evidence - ACCE

- The ACCE model covers evidence about the Analytic validity, Clinical validity, Clinical utility, and Ethical/legal/social implications (when applicable) to build the evidence base for the test.
- Analytic Validity: The technical performance of the test - does the test actually measure what it is supposed to. Common measures = sensitivity, specificity etc.
- Clinical Validity: The strength of the association between a genotype and disorder of interest. The strength of this association determines the test's ability to diagnose a disorder, assess susceptibility or risk, or provide information on prognosis or variation in drug response.
- Clinical Utility: Evidence that test results can change patient management decisions and improve net health outcomes

Method: Applying ACCE model

- Based on the EWG recommendations we developed
 1. An overarching question that sought to find direct evidence addressing our primary question
 2. A set of questions based on the ACCE that would build the chain of evidence that could help answer the overarching question indirectly

Methods : Overarching Key Question

KQ 1. Is there direct evidence that the addition of the specified molecular pathology tests used alone or in combination with traditional prognostic factors changes physician decision making and improves outcomes for adult patients with CRC, breast, lung, or bladder cancer compared with the use of traditional factors to predict risk of recurrence (RR) for adults with these cancers?

Ideally we hoped to find published evidence that directly answered this KQ.

Methods: Additional Key Questions

In the absence of direct evidence for KQ1, we developed KQs to build chain of evidence that would help answer KQ1. Chain of evidence based on ACCE model for evaluation Genetic Tests.

- **KQ 2. Analytic Validity:** Does existing evidence establish the technical accuracy and reliability of these tests for detecting the relevant molecular analytes?
- **KQ 3. Clinical Validity:** Does existing evidence establish the prognostic accuracy of the tests for predicting recurrence?
- **KQ 4. Clinical Utility:** Does existing evidence support clinical utility of the molecular pathology tests?

Methods: Additional Key Questions

Clinical Utility was further refined into impact on physician decision making and patient centered outcomes including harm.

- **KQ 4a.** What is the evidence that the prognostic information provided by the molecular pathology tests modifies physician decisions regarding use of adjuvant antineoplastic chemo- and/or radiotherapy, enhanced diagnostic testing for recurrence, and/or surgery among adult patients with malignant tumors?
- **KQ 4b.** What is the evidence that modified decisions lead to improved outcomes, including patient-centered outcomes such as improved quality of life, reduced disease recurrence, increased overall survival (OS) or disease-free survival (DFS), or reduced therapeutic side effects?
- **KQ 5.** What are the harms associated with treatment decisions that are informed by the molecular pathology tests?

Methods : Final Analytic Framework

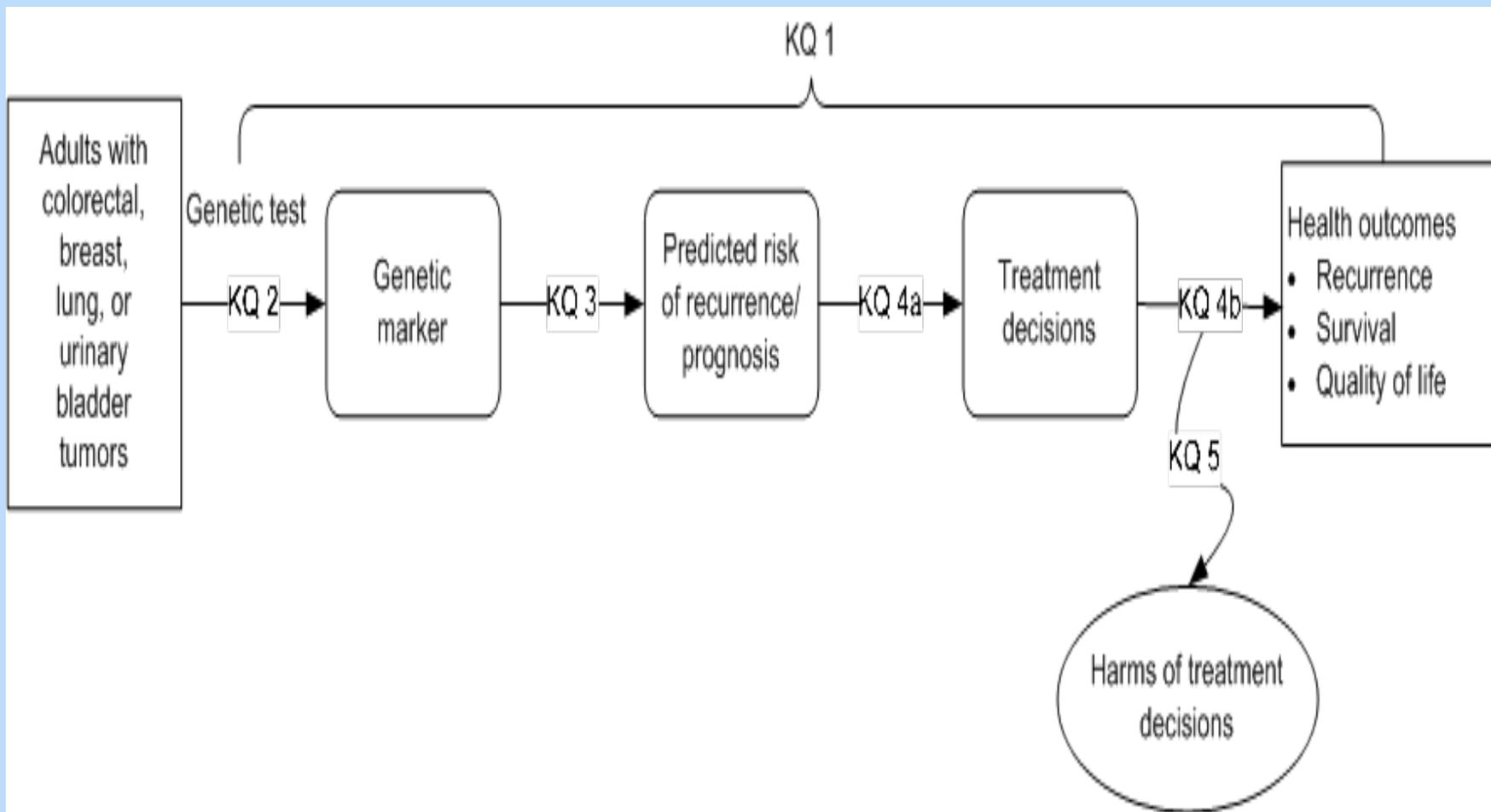


Figure A ES2

Methods: Searches

- PubMed[®], the Cochrane Library, and EMBASE[®] for English-language studies published through November 2013
- Also, searched reference lists of pertinent review articles and studies meeting our inclusion criteria
- Searched for unpublished studies relevant to this review using test developers' Web sites, ClinicalTrials.gov, the Food and Drug Administration Web site, Health Services Research Projects in Progress, and the European Union Clinical Trials Register
- Requested information from the College of American Pathologists (CAP) and from relevant companies, asking for data that they believe should be considered for the review

Method: Eligibility Criteria

- Population: Included studies of adult patients with one of the cancer types of interest that evaluated an eligible test.
- Intervention/Comparators: For KQs 1, 4, and 5, we included studies that compare at least 1 of the tests plus standard prognostic factors with the standard prognostic factors alone to determine whether the molecular pathology test adds independent prognostic value (benefit) or introduces additional harms (KQ 5).
- Did not include studies focused on patients with advanced/metastatic cancer or studies focused on predicting response to treatments.

Methods: Eligibility Criteria- PICOTS

- Comparators: For KQ 2 (analytic validity), we included studies of test performance, including intra/inter-lab reproducibility for included tests.
- Comparators: For KQ 3 (clinical validity), we included studies comparing patients with different test results (e.g., those with a mutation versus those who are wild-type) to establish prognostic value, with a multivariate analysis to adjust for known factors; we required that results were either adjusted for known prognostic factors or were specifically addressed in other ways, such as through inclusion/exclusion criteria of the study or stratification.

Methods: Study Selection and Data Extraction

- Independent dual review to assess for eligibility.
- Conflicts resolved by discussion.
- Used structured data extraction forms.
- One team member abstracted data; a second reviewed data for accuracy.

Methods: Risk of Bias Assessment

Assessed the risk of bias following the *Methods Guide for Medical Test Reviews* and *AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews* and the RTI Question Bank.

- For analytic validity used relevant questions from QUADAS-2 to assess potential for bias due to flaws in the sample selection, testing protocol, reference standards, verification procedures, interpretation, and analysis .
- For clinical validity and clinical utility, we assessed the potential for selection bias, confounding, performance bias, attrition bias, and detection bias.
- Two independent reviewers assessed each study.
- Assessed as Low, Medium, High, or Unclear.
- Conflicts between reviewers resolved by discussion until consensus.

Methods: Strength of Evidence

- Graded as high, moderate, low, or insufficient.
- Used the guidance established for the EPC Program.
- Incorporates four key domains: risk of bias (includes study design and aggregate quality), consistency, directness, and precision.
- Two reviewers assessed each domain for each key outcome and determined an overall grade based on domain ratings.
- Differences resolved by consensus discussion or by consulting with a third investigator.

Methods: Data Synthesis

For clinical validity (KQ 3), we conducted meta-analyses

- Estimated summary hazard ratios (HRs) for outcomes (for any given test-cancer pair) with three or more independent adjusted HR estimates.
- Tested the null hypothesis of homogeneity of effect sizes across the studies for each of the outcomes.
- If effect sizes non-homogeneous, summary effect size estimated using a random effects model.
- If effect sizes homogeneous, summary effect size estimated using a fixed effects model.

Results: Disposition of Articles

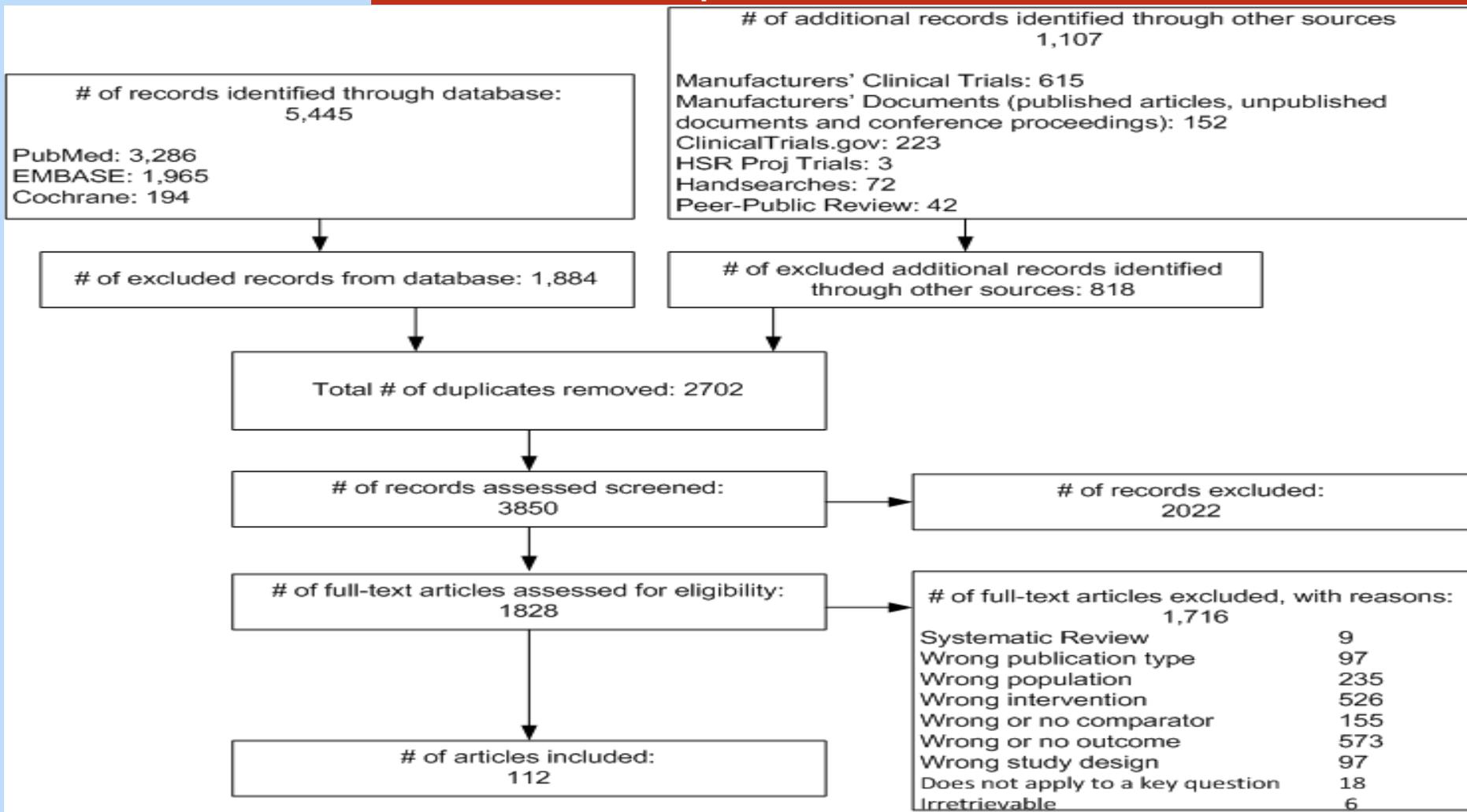


Figure B ES-5

Analytic validity

- Limited data on analytic validity in published literature.
- Published evidence was supplemented with proficiency tests results provided by the College of American Pathologists (CAP) for five tests. CAP focuses on inter-lab reproducibility.
- Based on CAP evidence BRAF, EGFR, KRAS, MSI and Urovysion are reported to have between 95 – 99% inter-lab reproducibility.
- Oncotype Dx is reported to have high intra-lab reproducibility by Genomic Health (Cronin, Clin Chem 2007).

Clinical Validity – Breast Cancer

Mammaprint: Poor prognosis vs. good prognosis

Evidence from multiple studies supports association between test result and prognosis for RR and CSS. Single study for OS.

Outcome	N Studies, N Subjects	HR (95% CI)
RR	6, 1,913	2.84 (2.11, 3.89)
CSS	5, 1,615	3.3 (2.22, 4.9)
OS	1, 144	1.67 (0.73, 3.82)

Clinical Validity – Breast Cancer Oncotype Dx: Comparing Hi-Risk to Lo risk

Evidence from multiple studies supports association between test result and prognosis for RR, CSS. Single study for OS.

Outcome	N Studies, N Subjects	HR (95% CI)
RR	7, 3,222	2.97 (2.19, 4.02)
CSS	2, 1,234	2.02 (1.35, 3.0)
OS	1, 668	1.65 (1.24, 2.19)

Clinical Validity –Lung Cancer

EGFR Mutation Testing: mutation vs WT

KRAS mutation testing: mutation vs WT

- Six studies looked at the prognostic value of EGFR for RR (n = 1,870) and OS(n= 1,820). Summarized evidence suggests no prognostic value.
- Some evidence that KRAS testing had prognostic value. Results displayed below.

Outcome	N Studies, N Subjects	HR (95% CI)
RR	4, 611	2.84 (1.14, 7.1)
CSS	0,0	
OS	2, 253	2.69 (1.19, 3.18) 3.33 (1.03, 10.82)

Clinical Validity – Colorectal Cancer BRAF Mutation Testing: WT versus Mutation

Evidence suggested added prognostic value of BRAF mutation testing in CRC for CSS and OS; Not significant for RR.

Outcome	N Studies, N Subjects	HR (95% CI)
RR	5, 4,106	1.07 (0.76 to 1.52)
CSS	7, 5,409	1.50 (1.26 to 1.77)
OS	11, 7,610	HR 1.45 (1.29 to 1.62)

Clinical Validity – Colorectal Cancer KRAS Mutation Testing: WT versus Mutation

Evidence suggested no added prognostic value of KRAS mutation testing in CRC for RR and OS ; added prognostic value for CSS.

Outcome	N Studies, N Subjects	HR (95% CI)
RR	5; 4,085	1.02 (0.91 to 1.14)
CSS	2; 1,174	1.30 (1.02 to 1.66)
OS	10; 5,328	1.22 (0.93 to 1.60)

Clinical Validity – Colorectal Cancer MSI Testing: MSI-H versus MSS

Evidence suggested added prognostic value of MSI mutation testing in CRC for RR, CSS and OS.

Outcome	N Studies, N Subjects	HR (95% CI)
RR	10; 7,130	0.60 (0.50 to 0.72)
CSS	6; 3,439	0.65 (0.51 to 0.82)
OS	12; 8,839	0.57 (0.43 to 0.77)

Clinical Validity – Colorectal Cancer Oncotype Dx

- One study with 690 patients reported results on the prognostic value of Oncotype Dx for RR on CRC.
- No published evidence that met our criteria for other outcomes.

Clinical Validity – Bladder Cancer Urovysion

- Urovysion was designed to be a diagnostic test for cancer of the bladder and not a test to assess prognosis.
- There is limited evidence (2 studies with a total n of 168) that it may be useful in predicting RR.
- No studies for CSS or OS.

Results: Clinical Utility: Patient Outcomes

- There were no published studies that assessed the impact of the test on long term outcomes for patients—e.g., impact on risk of recurrence or survival.
- Even in the cases where the tests seemed to add value in determining prognosis (i.e., evidence of clinical validity), we found no evidence that using the test was related to improved outcomes for patients.

Clinical Utility – Treatment Decisions

- Moderate evidence that Oncotype DX Breast, leads to changes in decisions.
- Although the decision changes were observed in both directions for individual patients, studies consistently showed an overall shift to less-intensive treatment recommendations as a result of using Oncotype DX Breast, with fewer recommendations for chemotherapy (and therefore less exposure to potential harms of chemotherapy).
- But studies did not follow patients to actually report on harms or to assess the overall balance of clinical benefits and harms.

Medicare Population

- No studies that focused solely on the Medicare population or assessed prognostic value of the tests stratifying for Medicare population.
- Almost all studies included patients from the Medicare population.
- We found no evidence to suggest that the clinical validity would differ for this population.

Limitations

- Many of the included studies had methodological limitations, introducing some risk of bias.
- For example, most of them were observational studies assessing associations between test results and outcomes, and are susceptible to potential confounding.
- No studies specific to Medicare population.
- Many of the included tests are currently used to predict response to specific treatments, an aspect that was not evaluated in this report.
- Determining whether the tests have clinical utility for predicting therapeutic response is beyond the scope of this review.

Summary

- The weight of published research to date has focused on the clinical validity of the tests of interest.
- Relatively little emphasis on how these tests can be integrated into the overall care of cancer patients in terms of changing decisions or the effect of those altered decisions on downstream patient-centered outcomes
 - Oncotype Dx Breast is the exception
 - With a relatively large number of studies showing an impact on treatment decisions resulting in fewer recommendations for chemotherapy (but still insufficient evidence on downstream outcomes)

Conclusions – Clinical Validity

Good evidence supporting added prognostic value (i.e., clinical validity), beyond traditional prognostic factors, for the following tests for RR, CSS, and/or OS:

- MammaPrint
- Oncotype DX Breast
- *KRAS* mutation testing for lung
- *BRAF* mutation testing for CRC
- *KRAS* mutation testing for CRC
- MSI for CRC

Conclusions – Clinical Utility

- Oncotype DX Breast leads to changes in treatment decisions, resulting in fewer recommendations for chemotherapy.
- No studies that directly assessed the impact of test use (for any of the included tests) on downstream health outcomes to establish clinical utility.

Thank you

Questions?