Technical Brief Disposition of Comments Report

Research Review Title: *Gene Expression Profiling for Predicting Outcomes in Stage II Colon Cancer*

Draft review available for public comment from May 7, 2012 through June 4, 2012.

Suggested citation: Black ER, Falzon L, Aronson N. Gene Expression Profiling for Predicting Outcomes in Stage II Colon Cancer. Technical Brief. No. 13. (Prepared by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-0058-I.) Rockville, MD: Agency for Healthcare Research and Quality. December 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Commentator & Affiliation	Section	Comment	Response
#2 Key Informant (KI)	General	This manuscript is very well-written and comprehensive. It will serve as an excellent primer on the state of current knowledge regarding GEP for stage II colon cancer.	No response needed
#5 KI	General	Overall, I think this evidence review is very thorough and did not have any major issues. I did wonder why at several times the idea of only testing T3 tumors without MSI was mentioned instead of suggesting that testing be restricted to all Stage II tumors without MSI but since it was just a theoretical subset, I don't think this matter.

In addition, I wonder if the language should be stronger that the summary is that use of these tests is premature at present until the gaps in the data chain can be filled. | No response needed

This is a summary of the evidence, so no changes were made from this comment |
#7 KI	General	The Technical Brief is complete, accurate, and very specific. The tables add a great deal to understanding it.	No response needed
#8 KI	General	I would request some discussion as to whether studies point to any single genes (I understand that they not within the scope of this report) that contribute to prognosis or prediction beyond the established risk factors such as MSI, histology, bowel perforation, etc.	This is beyond the scope of the report. The role of micro-satellite instability is recognized and that is included in the report. The literature for single-gene markers was not systematically-reviewed and is thus is not summarized in this report.
#10 Peer Reviewer (PR)	General	This report is well written and organized and I agree with the conclusions drawn	No response needed
#11 PR	General	Overall this is a well done review of a complex issue. The flow, key questions, review of the literature and analysis are well structured and thoughtfully presented. I have no substantive modifications.	No response needed
Public Agendia	General	For your consideration, we submit Nitsche, et al (J Onkologie, 2012), which includes an English language abstract. This German article publishes the validation series presented as an abstract by Rosenberg, et al, in 2010. This second validation study included only stage II colon cancer patients In addition, we performed a pooled analysis of stage II colon cancer patients from the validation studies 1 and 2 and additional cohorts from three hospitals. The pooled analysis presents 320 stage II colon cancer patients from five different hospitals in Spain, Italy, Austria and Germany. The pooled analysis was presented by Dr. Tabernero at the ASCO GI meeting 2012 and is presented at the poster discussion at the Annual ASCO meeting this week where he focuses on the subgroup analysis of T3/MSS patients. In each of the individual studies and in the pooled analyses, ColoPrint performance is compared with performance of clinical factors for risk assessment.	The Nitsche study (German with English abstract) was translated and information from this publication was added to the report; this information, along with information from Salazar, 2011, was added to Table 3 for prognostic markers..

Information from presentations at the 2012 ASCO meeting has been added to the report. |

Commentator & Affiliation	Section	Comment	Response
#4 PR	Abstract	Clinical problem (treatment of 100 stage 2 colon cancer patients with toxic chemotherapy to possibly benefit 3 or 4, and the inability to identify those 3 or 4 up front) clearly defined. The limitations of the existing studies were very well described. Most of the studies were for prognostic testing, but predictive testing is the more important clinical need. The studies do not demonstrate an ability of the GEPs to classify patients into risk groups to help conclude if adjuvant chemotherapy would be necessary/beneficial.	No response needed
		Bottom line very clear – no evidence exists to support the clinical utility of the GEP tests to alter or improve health outcomes and the risks/benefits of using the tests are not clearly defined. Much more evidence, both in quantity and quality, is needed (especially for predictive tests) before deploying these tests into practice; a good description of ongoing clinical trials is included in this report.	No response needed.
#9 Public Genomic Health	Background/General	It should be emphasized that conventional clinicopathological markers for stage II colon cancer decisionmaking have significant limitations. Although the authors acknowledge those limitations in the introduction, the rest of the document suggests that there is a failure to recognize that, if the same level of evidence standards were applied to existing markers (grade, LVI, perineural invasion, etc.), these markers would fare much worse than ODX on almost every front (analytical validity, clinical validity, etc.). Given what we know about traditional markers, particularly tumor grade, it is almost certain that patients are misclassified and, consequently, treated on the basis of an incorrect understanding of risk today. Moreover, with respect to absolute benefit, few existing markers have been rigorously examined for that relationship. One case in point is perineural invasion, which has, to our knowledge, only ever been studied in convenience cohorts, meaning that there are no data to support the notion that patients with perineural invasion should be treated because they derive larger absolute benefits, even though perineural invasion is included in existing NCCN© guidelines.	Thank you for the comment. This Technical Brief is a summary of the type of evidence available for the new GEP assays. It is not a review of current practice. No changes were made to the report.
Public Seidenfeld	Background	P 1 According to NCCN Guidelines Perhaps should cite AJCC staging manual rather than NCCN? I'm pretty sure that's where NCCN gets it from.	Citation has been changed to reference AJCC staging manual.
Public Seidenfeld	Background	P 1 Relative 5-year survival rate Since estimates for stage I and across all of stage II are provided earlier in the paragraph, it's probably worth adding estimates (ranges) here for OS @ 5 years for those diagnosed with stages IIA, IIB, IIC, III and IV post resection, to give a sense of how well stage (and stage subgroups) separates patients into prognostic subsets.	This would go beyond the focus of this report so no new information was added.

Commentator & Affiliation	Section	Comment	Response
Public Seidenfeld	Background	"Rate of response" or "response" is used throughout the document; however, in the adjuvant setting response per se is not measured, rather, "benefit" or relapse/disease free survival are commonly acceptable terms. Response is generally reserved for metastatic disease where actual tumor measurements are possible.	Agree with this comment. Throughout the manuscript, this has been changed to read "benefit from" adjuvant therapy. This comment was made several times.
#5 KI	Background	P 2 . . . patients with MSI-H tumors are excluded. Add the word tumors.	The word "tumors" was added
Public Seidenfeld	Background	P 2 Gene Expression Profiles/Predictive outcomes In adjuvant setting, may be better to talk about "effectiveness of" or "benefit from" than "response to" adjuvant therapy. If patients are NED, don't have a way to measure "response" so what you're really looking at is effect of adjuvant therapy on recurrence rate.	Agree, change made to read "benefit from treatment" (See note above)
Public Seidenfeld	Background	P2 Gene Expression Profiles/ GEP signature Need to distinguish between prediction of a general benefit from adjuvant therapy (across multiple effective regimens) versus prediction of benefit from a specific regimen or drug. These are separate clinical questions and probably should be addressed using different data sets or trial designs.	This section has been reworded for further clarity; report indicates that "the GEP result may only apply to regimen for which it was studied."
Public Seidenfeld	Background	P 2-3 Oral fluoropyrimidine Also combined with oxaliplatin in most patients (not entirely clear from the way this sentence is constructed). Also, some patients with T4 tumors that penetrate to a fixed structure may have added radiation therapy as part of the adjuvant regimen. But these are the highest-risk stage II patients, who likely would always receive adjuvant Tx. Perhaps they should be excluded from studies on clinical utility of GEP assays?	This has been reworded to specifically note "with or without oxaliplatin." Given the small number of patients where the question of radiation therapy may be considered, this was not added to the report.
#10 PR	Background	P2. In the first paragraph describing Gene Expression profiles, in the fifth line it states that, "Gene expression is determined by… DNA microarrays." The microarray data can be used to evaluate gene dosage. However, in the description of studies excluded from the literature review on page 11 the authors state that studies were excluded if "the testing did not analyze RNA". It seems that studies that include DNA microarrays could have or should have been included.	The GEPs analyze/measure RNA and this is done using the various techniques described; and use of DNA microarrays is one technique used for analyzing/measuring RNA. Thus, no changes are needed in the report.
#10 PR	Background	The point is well taken about the value of including test results describing MSI status of a tumor when considering GEP for prognostic or predictive information. Are there any other genetic test results that might be relevant to the validity or utility of GEP, such as KRAS mutation status? It might be helpful to briefly review these in the Background as well.	Nothing else (except MSI status) was noted in the review. KRAS was noted as a marker in advanced colon cancer.

Commentator & Affiliation	Section	Comment	Response
#10 PR	Background	The authors excluded studies of single gene expression profiling in colon tumors since that was not the focus of the report. However, it would be useful to include a description of these studies and the results regarding clinical validity and utility in the Background section of the report. It would be of interest to know if any of these single gene expression profiles that have proven validity or utility are also included in the multi-gene GEP testing.	Single gene studies are beyond the scope of this report which is on gene-expression profiles based on multiple genes. Since no systematic approach was taken to identifying single-gene studies, no summary information is presented.
#10 PR	Background	P.3 third paragraph describing Scope of Report. I recommend adding the phrase, "describing variants in or.." such that the first sentence would be changed to, "Because the topic of this Brief is GEP testing and GEP is based on results from expression of multiple genes, this Brief does not include reports describing variants in or expression of single genes."	Variants of genes seem to imply mutations which were also outside the scope and do not represent gene expression. Thus, statement added to report that studies of genetic mutations were also excluded from this Brief.
#10 PR	Background	P3. Third paragraph describing Scope of Report, in the last sentence it states that, "the use of GEP for other stages of colon cancer, or for both colon and rectal cancer, will not be included…" I think it would be important to state why rectal cancer is not included.	Information added noting that this is due to differences in surgical and adjuvant treatments for these two types of cancer.
#10 PR	Background	P3. Fourth paragraph, second sentence, I recommend adding the phrase "or non-responsiveness to chemotherapy". The statement should read, "Selecting appropriate patients who do not need adjuvant chemotherapy because of very low risk of recurrence or non-responsiveness to chemotherapy could improve the net health outcome by avoiding treatment-related adverse effects in those patients."	Agree. Change made.
Public Seidenfeld	Background	P 3 Their use should not be Agree; need to test prediction of clinical benefit separately for each specific regimen	No response needed
Public Seidenfeld	Background	P 3 The possibility of having similar findings Are there any relevant data on intratumor heterogeneity of colon tumors?	This is raised as a possibility, and an example is given of when this might be considered further. No changes were made.
Public Seidenfeld	Background	P 3 These single-gene markers Probably worth mentioning one or two examples as a parenthetical.	This is beyond the scope fo the report. See prior comments on this issue. No changes made.
Public Agendia	Background	P.3 Tumor Heterogeneity (comment) Please note that the impact of tumor heterogeneity is assessed for ColoPrint as part of the technical validation required for FDA submission.	FDA approval has not been given, so no changes made.

Commentator & Affiliation	Section	Comment	Response
#8 KI	Methods	This is a very well done comprehensive review. On page 5, under key informants, please add epidemiology as one of the disciplines of the clinical experts.	No response needed. Change made.
Public Seidenfeld	Methods	P 5 Manufacturer Should be plural (manufacturers)	Change made.
Public Seidenfeld	Methods	P 6 Patient-oriented and patient advocacy Perhaps this should have been the cancer.net site, not asco.org?	Information from the cancer.net site has been added to the report (Table 2)
Public Seidenfeld	Methods	P 6 Response to adjuvant chemotherapy How was response defined and measured in these resected stage II patients?	This has been changed throughout report to read "benefit from adjuvant therapy;" see comments above under Background section
#5 KI	Methods	P6 The data elements abstracted from included articles . . . The data elements which were abstracted from included articles . . .	No changes were made; the words "which were" were not added.
Public Agendia	Findings	P. 7 Table 1, The assay name is "ColoPrint® Colon Cancer Recurrence Assay" ColoPrint is commercially available worldwide to all stage II patients beginning on June 1st 2012. Specimen Used should indicate fresh tissue preserved in RNARetain. ColoPrint can be performed also on frozen tissue. The * footnote should indicate that ColoPrint is commercially available and may be available through PARSC Clinical Trial. (Supportive Literature: Press release "Agendia Announces Launch of ColoPrint for Colon Cancer Prognosis and Prediction," dated 01 June 2012; Agendia ColoPrint Specimen Sampling Instructions)	Table 1 updated to reflect this recent (6/1/12) change to note commercial availability.
Public Agendia	Findings	P. 7 Tissue Specimen, The ColoPrint test is performed on fresh tissue preserved in RNARetain. For logistic convenience, fresh tissue can be stored for up to 7 days in RNARetain and shipped at room temperature. ColoPrint can be performed also on frozen tissue. (Supportive Literature: Agendia ColoPrint Specimen Sampling Instructions). The RNARetain storage method has been cleared by FDA. http://www.accessdata.fda.gov/cdrh_docs/pdf7/K070675.pdf	Potential use of RNARetain has been noted in the report.
PublicSeidenfeld	Findings	P 7 GQ1 Currently Probably better to say "As of xx/xx/201x..." here, using date of last search.	Edited to read "As of July 2012, "which corresponds to submission of the final report.
PublicSeidenfeld	Findings	P 7 However, this issue This should be addressed in the section on analytic validity (last bullet of Question 3). Tissue handling and processing can have a major impact on analytic validity.	Aspects of handling that relate to GEP assays are discussed in the report; general issues that relate to laboratory tests in general are not discussed.

Commentator & Affiliation	Section	Comment	Response
#10 KI	Findings	P 7. (Beginning) There is a description of the GEPs. Please comment on the specific genes assessed in each assay, if possible, and describe the overlap between assays, if any. Also, as mentioned above, it would be of interest to know if any single gene expression profiles with evidence of clinical validity and utility are included in these multi-gene GEP tests.	Information about the overlap in genes (there was none) in three of the five commercially available GEP assays was added to the report. (Detail for genes in the other two assays was not found.)
#10 KI	Findings	P 7. Fourth paragraph, first sentence, there is a parenthesis after GEP that does not need to be there.	Edited.
#9 Public Genomic Health	Findings	There are many important differentiating factors which distinguish the Oncotype DX® Colon Cancer assay from the other tests which you reviewed. First and foremost, it must be emphasized that the evaluation of a test in a randomized clinical trial represents a quantum leap over studies in convenience cohorts. The fact that the Oncotype DX® Colon Cancer assay was validated in QUASAR, a landmark trial which randomized patients between observation and adjuvant 5FU/LV chemotherapy, enables Genomic Health, Inc. to make statements about the absolute benefit of adjuvant 5FU/LV being higher at high Recurrence Score® values than at low Recurrence Score® values, based on the evidence from the QUASAR validation study. No other test has similar data, and, thus, no other test can make that claim. Ultimately, an understanding of absolute benefit (and thus number-needed-to-treat) is what physicians and patients need to evaluate the risk-benefit trade-offs in the adjuvant decision making process. We feel that this should be included in the discussions on pages 7 and 11, when discussing Oncotype DX® in the context of other tests.	Items such as this will be very important as research is presented on the clinical utility of these measures. Information about the need for data from comparative studies related to predictive markers is discussed in the report. Data concerning the validation of the Oncotype DX® Colon Cancer assay with samples from the QUASAR study is included in the report; this study by Gray et.al. is reference 20.

Commentator & Affiliation	Section	Comment	Response
#9 Public Genomic Health	Findings	Finally, there are a few other areas which should be noted. There is an implication on page 7 that fresh frozen tissue may be preferable, due to less degraded RNA. The Oncotype DX® Colon Cancer assay addresses degradation with the use of reference genes, which provide an internal control for sources of pre-analytical variability which can impact on RNA degradation (e.g. type of fixative, duration of fixation, etc.). Further, there are ample data demonstrating reproducibility and precision of our RT-PCR platform from formalin-fixed paraffin-embedded tissue (FFPE). For fair balance, we would note that there are acknowledged limitations of frozen tissue in terms of proper tissue handling and for accurate microdissection. This data, as well as information on our sample turn-around time, which was listed as unknown, are available on our website, www.oncotypedx.com. The turn-around time for the Oncotype DX® Colon Cancer assay is 10-12 days, with 1-2 days more for MMR testing. One last suggestion is the addition into Table 9 of a recently closed 200 patient clinical utility study with the Mayo Clinic Cancer Research Consortium (http://mccrc.mayo.edu/trials/control/colorectalcolon.html), evaluating realworld impact on decision-making in the treatment of stage II colon cancer. The Mayo study is expected to be reported in late 2012.	Section edited to clarify that there are two methods currently used for GEP assays. Brief comments made about potential pros and cons of methods.

Information added to report.

Only data available from peer-reviewed publications were included in this Brief. |
#3 PR	Findings	There is one typographical error on page 8, line 56 which reads "This is use as a predictive..." and should read "This is used as a predictive..."	This change has been made.
Public Seidenfeld	Findings	P 9 20 to 25 percent reduction Doesn't this markedly overstate the magnitude of benefit from adjuvant therapy in stage II patients?	This section has been reworded and this phrase has been removed. The original phrase should probably have indicated "20 to 25 percent relative (not absolute) reduction" to avoid confusion.
Public Seidenfeld	Findings	P 9 Variability in clinical studies Shouldn't this quote or cite specific studies that did versus did not limit inclusion to stage II patients? Similarly, those that did versus did not limit to those without MSI?	Since this is a Technical Brief have not added this additional detail. However, information about MSI is included in Table 3.
Public Seidenfeld	Findings	P 10 GEPs marketed as LDTs So for the 3 assays listed in Table 1 as available outside of clinical trials, are they being developed for marketing as LDTs or as test kits? Not clear to me from what said in this paragraph.	They are LDTs. Report edited to reflect this. (Note that with the update Table 1 now has 5 assays.)

Commentator & Affiliation	Section	Comment	Response
#10 PR	Findings	On pages 9 and 10 the authors describe the FDA status of these tests. In the last sentence in the paragraph ending at the top of page 10 it states, "GEPS marketed as LDTs may enter the U.S. market with analytical validation under laboratory regulations imposed by CLIA but without evidence of clinical validity or utility." This is a true statement, but the way this is written it may lead the reader to think that that having FDA oversight would require evidence of clinical validity or utility. To my knowledge, this is not the case for LDTs. Can you comment on this? Is it true that neither CLIA nor FDA require evidence of clinical validity or utility for approval?	Sections on FDA have been reworded to clarify the issue. Comment added that FDA approval would likely require data on at least clinical validity, but not necessarily clinical utility.
#10 PR	Findings	P. 10 (Bottom) Last line, the number 1,110 is separated with 1, on page 10 and 110 on top of page 11.	Now fixed/resolved.
#11 PR	Background	P. __Sentence on lines 32-34 starting with "Selecting appropriate patients" is awkward.	Sentence has been reworded. Now reads, "Identifying patients who do not need adjuvant chemotherapy because of very low risk of recurrence or because of predicted lack of benefit from adjuvant chemotherapy would improve the net health outcome by avoiding treatment-related adverse effects."
Public Seidenfeld	Findings	P 11 202 full-text Need a noun here. e.g., full-text articles Also, might be worth adding a QUORUM diagram here.	Change made (noun added)
Public Seidenfeld	Findings	P 11 None of the studies I'm uncertain that clinical utility is a relevant question for a purely prognostic marker. What's more to the point is whether it adds independent prognostic information to the clinical factors used to "guesstimate" prognosis without the marker. But I agree that reclassification analyses and ability to discriminate discrete prognostic groups are key points, and it's problematic (and rather surprising) if that really is absent from all published studies. Not even separate survival curves for high, intermediate and low risk groups as defined by the GEP versus the group as a whole???	Prognostic markers could have clinical utility if they can exclude patients from receiving adjuvant therapy. Detail about this added to the report. Publications do show separate curves for risk groups, but this does not show clinical utility.

Commentator & Affiliation	Section	Comment	Response
Public Seidenfeld	Findings	P 11 None of the publications I don't agree with this statement. The Oncotype DX validation study by Gray et al (2011; ref. 19) looked at survival curves with versus without adjuvant therapy separately for low, intermediate, and high RS groups. Admittedly, I did not think the assay SUCCESSFULLY or USEFULLY separated groups by likelihood of benefit from the regimen used in the RCT that provided samples, but at least they attempted that kind of analysis.	This was meant to show successful validation – and as noted the validation was not successful. Notation (abbreviation) on Table changed so that NV means "not successfully validated" The lack of successful validation was/is noted in the text of the Brief.
Public Seidenfeld	Findings	P 11 GEP was derived This should be pleural, not singular: the GEPs were derived on groups of colon cancer patients.	Edited to read "The GEPs were derived …."
Public Seidenfeld	Findings	P 11 In 2011, Grone I'm not sure how this connects with the rest of this paragraph. It's important to include mention of the inability to replicate the reported results, but perhaps better to break into separate paragraphs?	Made this a separate paragraph.
#10 PR	Findings	P. 11 last paragraph, the second sentence should probably be changed to "The GEP tests were derived from tumors from patients ranging in number from 22 to 1,851 (median=55)."	Sentence reworded.
Public Agendia	Findings	P. 11 MSI-status, Note that all published clinical validation studies of ColoPrint take MSI status of patients into account. All studies show that ColoPrint is independent of MSI status. Most patients with MSI-High status are correctly identified as low risk by ColoPrint. Also, note the analysis in the ASCO posters that describe ColoPrint performance in the T3/ MSS subgroup of stage II patients.(Supportive Literature: Salazar, Pooled Analysis, ASCO 2012 Tabernero, Pooled Analysis, ASCO GI 2012 Salazar, 2011)	Table notes that the ColoPrint assay does take MSI into account.
#5 KI	Findings	P 11 Review of the 202 full-text identified . . . Review of the 202 full-text articles identified. A list of data elements abstracted . . . A list of data elements which were abstracted . . .	Changed Okay as is.

Commentator & Affiliation	Section	Comment	Response
Public Agenda	Findings	P. 12 Reclassification, Please note that all published clinical validation studies of ColoPrint compare ColoPrint performance with risk assessment using the current clinical ASCO and/ or NCCN guidelines. The results demonstrate that ColoPrint is better than clinical factors and MSI status to separate low and high risk groups but also indicate that a combination of ColoPrint with clinical factors might even further improve risk classification.(Supportive Literature: Salazar, Pooled Analysis, ASCO 2012)	Information from the net reclassification in the derivation study of Salazar, 2011 has been added to the report. As noted in a response to a previous comment, the Salazar, 2011 paper was added to Table 3. The Brief does not summarize information from meeting abstracts, the data in the Brief is based on peer-reviewed publications.
Public Seidenfeld	Findings	P 12 An eleven-gene Table 3 identifies this GEP as the Oncotype DX Colon Cancer Assay; Table 5 probably should do so as well, since the same two references (#'s 19 and 25) are cited.	The 11 gene predictive assay is different from the 12 gene prognostic assay; the word "unique" added to clarify that this is a different marker even though the studies (publications) are they same. At present, only the 12 gene prognostic assay is referred to as the Oncotype DX® Colon Cancer Assay.
#9 Public Genomic Health	Findings	P 13, Table 3 The evaluation of patients on clinical trials (as has been the case for Oncotype DX®) allows for the study of patients enrolled with clear inclusion and exclusion criteria, unlike convenience cohorts where such criteria are not applied and multiple biases (particularly choice of treatment) must be considered. The interpretability of study results must account for these differences. The other tests described in your review have been based on studies in convenience samples, and have not been studied in clinical trials, to our knowledge. We believe this distinction should be noted. Furthermore, the size of our studies and our focus on stage II in validation should not be overlooked. The confidence that comes with consistent results across large, well-powered studies is what is required for physicians and patients to make important clinical decisions in practice. This advantage for ODX is starkly apparent in Table 3. In addition, it appears that the authors are not aware of the 2010 paper by Clark-Langone, et al 1, which demonstrates the analytical performance of the Oncotype DX® Colon Cancer assay, including reproducibility of test results.	The report shows that the Oncotype DX assay has been successfully validated and size of the sample was noted. This information has been added to the report and to Table 4.

Commentator & Affiliation	Section	Comment	Response
#11 PR	Findings	P. 15 Line 46, the "(very)" in parenthesis seems equivocating without purpose. Either "very" should or should not be in the sentence.	Term "very" removed.
#11 PR	Findings	P. 16 Sentence starting on line 28, would read better if the first "possibility" is struck from the sentence.	Agree, sentence edited and "possibility" removed
#1 PR	Findings	On page 16 under guiding question 2 line 37[sic] (p 21, guiding question 4?), a very critical group that may or may not be included in any of the publications concerning stage II patients is the group that has fewer than 10 or 12 nodes reported. This has been a moving target over the last 5 years as more patients have the requisite 12 nodes reported. However, the retrospective studies may exclude those with fewer than 12 or 10 nodes or they may include everyone. This would have a great impact on outcomes and performance of any prognosticator due to the stage migration phenomenon in the adequately sampled patients. This issue should be included and considered in the first paragraph under this question as well as in the general analysis.	Additional detail added to report about this topic. Report now describes nodes "assessed" and indicates this is now an important quality marker for colon cancer surgery.
Public Agendia	Findings	P. 18 Abstracts … for potential breaking scientific developments (response to request in Draft Brief for additional detail): A summary of the Rosenberg (2010) abstract including 135 stage II colon cancer patients requested additional details about this study. Nitsche, et al, J Onkologie, 2012 (English language abstract included) reported on the n=135 stage II patients included in this validation study from the total cohort (n=233) stage II and III colon cancer patients. Nitsche reported that ColoPrint identified most stage II patients (73%) as low risk. The 5-year distant-metastasis free survival was 95% for low risk patients and 80% for high risk patients.	Information from the Nitsche article (translated from German) has been added to report.
Public Agendia	Findings	P. 19 Abstracts … for potential breaking scientific developments (response to request in Draft Brief for additional detail) The pooled analyses described by Tabernero and Salazar include only stage II colon cancer patients from five hospitals in four European countries (Germany, Spain, Austria, Italy). The pooled analysis includes cohorts from the validation studies 1 and 2 that have been previously reported by Salazar from Spain (n=103) and by Maak/Nitsche/Rosenberg from Germany (n=135). In addition, the pooled analysis has patients from Vall D'Hebron (Barcelona), University of Vienna Hospital (Austria) and Ferrara (Italy). Tabernero reported "ColoPrint classifies two-thirds of the stage II patients as Low Risk. The 3-year RFS was 91% for Low Risk and 74% for High Risk patients with a HR of 2.9 (p=0.001)." (Supportive Literature: Tabernero, Pooled Analysis, ASCO GI 2012; Salazar, Pooled Analysis, ASCO 2012)	Summary of the 2012 ASCO Annual meeting (06/12) abstracts added to report.
Public Seidenfeld	Findings	P 19 In 2012 Tabernero Just as an FYI, the May 1 issue of ASCO Post published a brief summary of data from this study that were presented at ASCO's GI Cancers Symposium.	ASCO Abstract in report.

Commentator & Affiliation	Section	Comment	Response
Public Agenda	Summary	P. 21 Inadequate number of Lymph Node (LN) assessment, Please note that in all ColoPrint studies to date, a very low percentage of patients had less than 12 LN assessed (20% in the pooled analysis). (Supportive Literature: Tabernero, Pooled Analysis, ASCOGI 2012).	Issue of lymph node assessment has been emphasized in report.
#1 PR	Summary	Under guiding question 4 on page 21 line 31 through 36 there is a discussion concerning deficiencies related to a lack of reclassification analysis and risk categorization in the available studies. I can't speak to each of the studies but the OncotypeDX publications have very clearly segregated patients into risk strata. This group and others have also shown that their particular prognosticator is independent of other commonly used metrics of risk. I'm not sure it is fair to say that the publications have not addressed the risk stratification issue although it is definitely arguable that the clinical value of the stratification may not be very high.	Results that segregate (classify) patients provide information about clinical validity. However, as noted in the Brief, to demonstrate clinical utility data must be provided to show how use of the GEP assay in clinical care impacts net health outcome. The results noted in the comment do not permit conclusions about impact on net health outcome.
#1 PR	Summary	Page 21 line 52, although sampling is one reason for low LN counts, there are other reasons including number actually examined and/or reported by the pathologist. This is a sensitive issue with surgeons and pathologists.	Detail added that this can be due to sampling at surgery or extent of pathological review
#5 KI	Summary	P 21 . . . alone, bur . . . alone, but	Corrected
Public Agenda	Summary	P. 22 Table 10 (correction) ColoPrint has been validated for recurrence in colon cancer patients only. ColoPrint has been validated for reclassification – about 45% of patients have discordant clinical and ColoPrint risk assessments.	As noted in prior responses, information about ColoPrint assay was added to Table 3 based on Salazar, 2011 and Nitsche, 2012. This added information was used to update Table 10 which now shows ColoPrint validated for recurrence in stage II colon cancer and shows the net reclassification from the derivation report of Salazar, 2011.

Commentator & Affiliation	Section	Comment	Response
Public Seidenfeld	Summary	P 22 Table 10 Oncotype DX Again, I think this may not be correct. The paper by Gray et al. (2011) did include an analysis to address effects of FU/LV on recurrence rates separately in the three risk groups. The results were far from impressive, but at least they tried.	The predictive assay (eleven gene expression treatment score) was not validated as reported in the publication by Gray et al. The 12 gene prognostic assay (Oncotype Dx® Colon Cancer assay) was not being studied to predict benefit from adjuvant chemotherapy. This is the reason for the entry of "No."
Public Agendia	Summary	P. 23 Box 1 (comments and additions) Clinical Validity: ColoPrint has been validated in Spanish, Northern European and US patients. The PARSC study includes also Asian patients. Patients: ColoPrint has been validated in all stage II patients and in patients with T3/MSS phenotype. ColoPrint is repeatedly the only factor that is significant in the multivariate analysis, indicating that it is independent and superior to other factors. Clinical Utility: ColoPrint has been compared to clinical risk stratification. ColoPrint better distinguishes low and high risk groups as compared to risk assessment using current NCCN guidelines. This is partially due to the fact that ColoPrint identifies more patients as low risk (~65%) than the clinical factors and therefore is in better agreement with clinical reality that shows nearly 75% of patients have no relapse after 5 years. However, since ColoPrint and clinical factors recognize different subgroups (as indicated by the 45% discordance in risk classification of clinical factors alone), the combination of ColoPrint and clinical factors gives better risk stratification.	The brief contains data on ColoPrint from peer-reviewed publications, and it also summarizes abstracts from recent scientific meetings. Detail was added to the Brief to clarify what data are needed to show clinical utility; that is, whether use of the GEP assay in clinical care improves net health outcome. The report indicates that clinical utility has not been shown with data from peer-reviewed publications. Information in studies from meeting abstracts needs greater detail (through peer-reviewed publication) to determine whether any conclusions can be made about clinical utility.

Commentator & Affiliation	Section	Comment	Response
#10 PR	Summary	P 23. In Box 1, I believe the questions under the topic, Patients, belong with the questions relating to the Clinical Validity topic. In addition, the last question listed under the Clinical Utility topic would also seem to be more appropriate under the Clinical Validity topic. It reads, "How does use of GEP compare with use of other approaches to risk stratification (classification)? The latter change would also be consistent with wording on the following page (24).	While the items related to "Patients" could have been included under clinical validity, this was kept as a separate heading so that the importance of these questions could be noted. Clarification of "patients for testing" will be very important as clinical algorithms for use of GEP assays are developed and tested. The comment about comparison of GEP with other approaches is considered most relevant to clinical utility; thus its current placement. While this can be considered with clinical validity, the key question is how use of the GEP assay in clinical care impacts net health outcome (clinical utility) compared to other approaches.
Public Seidenfeld	Summary	P 23 As shown in the evidence map Is it really necessary to repeat all this?	This section has been edited to reduce repetition.
Public Seidenfeld	Summary	P 23 Box 1 (response to therapy) Again, what's meant by "response to therapy" in a patient with resected stage II disease? How is "response" defined and measured?	As noted earlier, changed to "benefit from" or improved survival and/or decreased recurrence through report.
Public Agendia	Summary	P. 24 (as stated above) Clinical Utility: ColoPrint has been compared to clinical risk stratification. ColoPrint better distinguishes low and high risk groups as compared to risk assessment using current NCCN guidelines. This is partially due to the fact that ColoPrint identifies more patients as low risk (~65%) than the clinical factors and therefore is in better agreement with clinical reality that shows nearly 75% of patients have no relapse after 5 years. However, since ColoPrint and clinical factors recognize different subgroups (as indicated by the 45% discordance in risk classification of clinical factors alone), the combination of ColoPrint and clinical factors gives better risk stratification.	A detailed, net reclassification analysis (how patients in each risk group are impacted by use of the GEP result) from large validation studies along with analysis showing the impact of this net reclassification on net health outcome is needed in order to demonstrate clinical utility.
#3 PR	Summary	The references to Figure 1 on research questions (page 24) are a little difficult to follow since research questions 4 and 5 (shown in the figure) are not addressed in the text (page 24).	Have tried to address this; have included question 5 as part of clinical utility (understanding overall benefits and harms).

Commentator & Affiliation	Section	Comment	Response
#10 PR	Summary	P. 24 In the second bullet it states, "Determining the clinical validity (accuracy) of the GEP assays in classifying patients…."	This "Next Steps" section has been modified and a sub-section has been added on important questions for future research. The note on Clinical Validity now reads "How strongly is the GEP assay result correlated with the outcome of interest?"
#10 PR	Summary	P. 24 In the second bullet, I would create a new bullet beginning with the statement, "To date, there has been much less research on use of GEP assays as predictive (response to therapy) markers…" To me this topic relates to clinical utility and not clinical validity, which seems to be the focus of the second bullet.	This section has been extensively edited.
#10 PR	Summary	Perhaps another important question for future research would relate to identification of patient, provider and organizational factors contributing to successful implementation of GEP testing for prognosis or response to treatment.	The scope did not consider implementation, so this issue was not noted as a research need.
#11 PR	Summary	Stated as P. 28. I think the construct presented at the bottom of page 28 in the Guiding Question 4 section, might be better presented. I think the core concept is that it is uncertain but unlikely, that GEP testing will be able to predict recurrence when the recurrence is related to surgical technique. I think this can be stated in a more straight forward manner.	Edited to read that "one could assume that results of a GEP assay would not predict future events related to the surgical procedure, e.g., a recurrence related to a resection margin that showed cancer cells."