

Project Name: **In-Situ Fluorescent Hybridization (FISH) or Other In Situ Hybridization (ISH) Testing for Chromosomal Damage in Uterine Cervical Cells to Predict Potential for Dysplasia/Malignancy**

Project ID: **CANC0511**

Table 1: Invited Peer Reviewer Comments

Reviewer <sup>1</sup>	Section <sup>2</sup>	Reviewer Comments	Author Response <sup>3</sup>
1	General	The review of the basic facts regarding HPV and cervical carcinogenesis is not expert. There are multiple minor errors that hurt the credibility of the report.	The background was revised. We replaced the outdated terms: dysplasia and CIS. We revised the disease progression model. We introduced the LAST terminology for histopathological squamous intraepithelial lesions.
2	General	Overall, the results of the review are consistent with my own independent review and knowledge of this area.	Thank you.
3	General	I agree with the major conclusions of this study. In particular, the available evidence does not suggest that ISH testing has superior sensitivity or specificity compared to existing testing. The conclusions of this study are based on a literature review to July 2012, and I am unaware of any new data that contradicts your basic conclusions. Though I have not performed a formal literature search for very recent data, there are two abstracts to be presented at the American Society of Cytopathology meeting in November 2012 that studied FISH testing (3q26) as an adjunct to Pap testing, and neither study showed value of testing.	Thank you.  As of 2.27.2013, these abstracts have not been published.
3	General	I also double-checked two key articles against the data in your Technology Assessment. The first has been cited to favor the utility of FISH testing: Jalali et al., <i>Amplification of the chromosome 3q26 region shows high negative predictive value for nonmalignant</i>	Thank you.

		<p><i>transformation of LSIL cytologic finding.</i> Am J Obstet Gynecol 2010, 202(6):581-85. I agree entirely with your assessment that the findings in this article are not compelling and the study design has potential for bias. I also reviewed an article that you excluded from analysis due to insufficient ability to correlate the FISH results with histologic biopsy: Caraway NP et al. <i>Gain of the 3q26 region in cervicovaginal liquid-based pap preparations is associated with squamous intraepithelial lesions and squamous cell carcinoma.</i> Gynecol Oncol 2008, 110(1):37-42. In spite of the title of this paper, I agree that in fact you are correct that this article really does not have enough correlative data to be useful.</p>	
3	General	<p>Starting on page 34, the document repeats itself, though with more detail. These two sections need to be merged, and references updated between them. I could re-review this Assessment when the merging of the text is completed.</p>	<p>The 34<sup>th</sup> page of the document marks the beginning of the main report. The prior 34 pages contain the abstract and executive summary. Executive summaries provide a concise summary of the report.</p>
3	General	<p>It may be useful to compare some other triage methods to the current meta-analysis. For example, The Aptima HPV E6/E7 mRNA test had sensitivity of 96.3% and specificity of 43.2% in a high prevalence cohort, and a specificity of 88.3% in a low prevalence screening cohort. (JOURNAL OF CLINICAL MICROBIOLOGY, Feb. 2011, p. 557–564 Vol. 49, No. 2). This comparison is useful to interpret the utility of your meta-analysis in which you conclude:</p> <p>“Five studies compared the clinical validity of TERC in LSIL for CIN3+, with two testing FISH for TERC or MYC. Again, only one study tested patients who were positive for HPV. In these studies, the sensitivity ranged from 0.45 to 0.93 and the specificity ranged from 0.42 to 1.00. Meta-analysis of five studies of TERC in LSIL for CIN3+ found summary sensitivity of 0.78 (95 percent CI 0.65, 0.87) and summary specificity of 0.79 (95 percent CI 0.51, 0.93).”</p>	<p>We referenced the Aptima. 2011study in the discussion to provide some context.</p>

4	General	<p>Overall, this is a relatively complete report with no novel findings since it the lack of validity of these FISH assays is already known in the scientific community. However, FISH apparently is being used in some pathology laboratories, hence, giving the review seems worthwhile.</p>	Thank you.
5	General	<p>I was surprised to see that a Healthcare Technology Assessment Report was commissioned for this topic. The studies evaluating ISH for cervical cancer screening are mostly small, using convenience samples, very different assays, and provide very heterogeneous results. Is there an “evidence threshold” to move forward with a full HTA review?</p> <p>I was also surprised that a major focus of the report was to evaluate the role of ISH assays for Medicare populations, since the typical age range of women with Medicare is outside the window of cervical cancer screening.</p>	<p>As outlined by the comment by reviewer 4 (above) FISH tests are being marketed and offered by some laboratories, thus a review was commissioned to assess the technology.</p> <p>We added this sentence in the overview of the Executive Summary and in the Background to the main report: “Tests being marketed and offered by some laboratories include in situ hybridization (ISH) tests, including fluorescence ISH (FISH), to detect HPV or chromosomal abnormalities.”</p> <p>While the majority of the population covered by Medicare is over 65 years, about 15% are younger than 65. Thus this report is of interest to Medicare. Further, Medicare coverage decisions are often followed by other payers.</p>
5	General	<p>The review has been conducted very thoroughly and represents the current evidence of ISH assays for cervical cancer screening very well.</p> <p>Still, the conclusions should be more critical of the available data and point out clearly which studies are needed to address the</p>	<p>Thank you.</p> <p>We believe our conclusions are critical in calling the evidence insufficient and immature. We also added (additions in bold) to the conclusion: <b>Although ISH tests for cervical cancer testing are marketed and offered by some laboratories</b>, limitations of the evidence base are the lack of <b>evaluation of ISH in</b></p>

		questions properly.	<p><b>current</b> screening contexts, the <b>lack of test standards, etc</b></p> <p>We added a suggestion by the public reviewer (see below) in the research gaps section of the discussion “An expert conference may be helpful to agree on common measurement guidelines.”</p> <p>We further included the recommendation to use LAST nomenclature to categorize histology outcomes going forward.</p> <p>The research gaps section also includes other specific recommendations on evaluating the value of ISH testing in the setting of co-testing with PAP and HPV, and in the setting of upfront HPV genotype testing, the need to investigate panels of probes to enhance diagnostic accuracy, the need to use modeling to project the impact of testing on clinical management, and the need to evaluate ISH testing for detection of adenocarcinoma.</p>
6	General	A well performed review of available literature. Agree with conclusions that were made.	Thank you.
4	Title	The title somehow implies that the only reviewed FISH or ISH is those that test for chromosomal damage. Whereas HPV 16 and 18 were chosen along with TERC and MYC. The rationale for including HPV 16 and 18 is not clear or the title should be changed.	<p>The reference to chromosomal damage was removed from the title. It now reads:</p> <p>Fluorescence In Situ Hybridization (FISH) or Other In Situ Hybridization (ISH) Testing <del>for Chromosomal Damage</del> in Uterine Cervical Cells to Predict <del>Potential for Dysplasia/Malignancy</del> Precancer and Cancer</p>

1	Abstract	It is “high-grade” not “high-risk” lesions that constitute precancer.	We have changed this throughout when referring to CIN lesions
5	Abstract	Clarify that a number of the ISH assays that were evaluated test for multiple types, not just HPV16/18	We changed this to plural in the abstract.
5	Abstract	Comment on cytology is unclear: Why were only ASC-US/LSIL results included? One possible application, as outlined by the authors, would be triage of HPV+/Cyto- women. Thus, evaluating normal cytology specimens seems important	We added a section stating that we found no relevant data in groups with HPV+/Cyto- samples.
5	Abstract	It is unclear why a tradeoff between sensitivity and specificity suggests a threshold effect. Continuous biomarkers typically show a tradeoff between sens/spec when the cutoff is moved. In contrast, a threshold effect describes a qualitative change at a certain threshold.	We used “threshold effect” to denote the negative correlation between sensitivity and specificity when studies have different explicit thresholds for “positive” tests. As the reviewer suggests, in a single study sensitivity and specificity are inversely related. Similarly, across studies, one expects that, if a threshold effect exists, studies with a higher estimated sensitivity will have a lower specificity (after accounting for heterogeneity and random variation [Trikalinos TA, J Gen Intern Med › v.27(Suppl 1); Jun 2012]). We agree that there are other possible causes that can explain this pattern of results, so we changed the text to: which <b>may be indicative</b> of a threshold effect”
5	Abstract	A major issue related to ISH assays is the evaluation of samples. This topic was not really addressed in the review. Scoring of ISH slides can be very time-consuming. Automated approaches are promising, but data are lacking. In order for ISH to become a routine test, the evaluation of test results need to be standardized and accelerated.	We agree. We have emphasized the need for test standardization and acceleration in the sections on research gaps and conclusions.

1	Executive Summary	Dysplasia and CIS are outdated terms; I suggest not using this old progression model. The newer model is that HPV infection (low-grade lesions or CIN1) if persistent can progress to precancer (high-grade lesions or CIN2-3 roughly, although not all CIN2 qualifies). The latest revisions of histopathology do not mention dysplasia or CIS at all.	Dysplasia and CIN were replaced in the title and text except when the articles we reviewed explicitly referred to the former terms.  The newer disease model is also now described and the recently published LAST terminology has been introduced.
1	Executive Summary	To say there are physical harms associated with colposcopy (even including biopsies) is really stretching the reality. Does TOMBOLA really indicate that?	Tombola concluded: "Cervical punch biopsies and, especially, LLETZ carry a substantial risk of after-effects."  We changed "psychological and physical harms" to "adverse events."
1	Executive Summary	Histology does not provide definitive diagnosis. Histology has errors, including errors from colposcopy that is quite imperfect. For example, an HSIL-CIN3 cytologic interpretation, even if biopsy is CIN1, is a high-risk state.	This paragraph has been edited.
1	Executive Summary	Integration of HPV DNA into human genome is NOT required for cervical cancer. It is an association, not a necessity.	This sentence has been removed.
1	Executive Summary	The high risk or carcinogenic HPV types are best defined by IARC, and include HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59. HPV68 is likely carcinogenic.	We summarized the HPV types that are implicated most commonly in cervical cancer according to IARC. The list you have provided are the most prevalent types which we considered different in its impact on cancer.  We were able to collect this statement from the IARC report: "Worldwide, HPV-16 and 18, the two vaccine-preventable types, contribute to over 70% of all cervical cancer cases, between 41%-67% of high-grade cervical lesions and 16-32% of low-grade cervical lesions. After HPV-16/18, the six most common HPV types are the same in all world

			regions, namely 31, 33, 35, 45, 52 and 58; these account for an additional 20% of cervical cancers worldwide (Clifford G et al. Vaccine 2006;24(S3):26-34).”
1	Executive Summary	SurePath should also be referenced for fairness, along with ThinPrep. Why choose just one?	SurePath was added.
1	Executive Summary	The phrase “before cellular abnormalities are evident” is not clear.	This was removed from the sentence.
1	Executive Summary	If FISH is one example of ISH, what are other detection systems? Describe briefly.	The section was expanded to include other detection systems.
1	Executive Summary	The sine qua non of a successful molecular diagnostic is reproducibility. The lack of data on this topic is a major deficiency in the evidence base. In my mind, it is enough, and one could stop and say at this point that the method is not ready for clinical use.	We have emphasized the need for standardizing measurement in research gaps and in the conclusion.
1	Executive Summary	Do the authors really mean that the data are spurious?	This sentence has been edited and now reads: “Overall, for questions related to preanalytic issues impacting analytic validity, the data were sparse and highlighted a lack of commonly agreed upon test and validation standards.”
1	Executive Summary	Which clinical guidelines specifically say that ISH can be considered for women ...[Key Question 3]? I don’t recall the 2012 ACS/ASCCP/ASCP guidelines recommending this.	We have changed this to say: “ISH may be considered...”
1	Executive Summary	Is meta-analysis valid when the studies vary so much, and are so heterogeneous? If so, sensitivity is low for CIN3+.	We agree that the data are statistically heterogeneous. This heterogeneity is taken into account by the random effects model we used (resulting in “appropriately” wide confidence intervals for the summary estimates of sensitivity and specificity). However, we do not base the decision to perform meta-analysis on observed

			statistical heterogeneity. The studies we synthesized were conducted in women with ASCUS or LSIL, using FISH as the index test with a histological diagnosis of CIN2+ or CIN3+ on biopsy. On the basis of these similarities we judged that the studies could be quantitatively synthesized.
1	Executive Summary	This phrasing was not clear to me: “The evidence was considered to be direct for clinical validity. Thus, overall we have low confidence that the estimated clinical validity of the FISH test represents its true validity.”	We added the following sentence to clarify: “The evidence was considered to be direct for clinical validity since the studies examined CIN2 or 3, which were considered important outcomes for decisionmaking.”
1	Executive Summary	The Executive Summary for key question 3b. does not seem to be directly relevant to the ISH topic.  KQ 3b relates to: How similar are the spectrum and prevalence of the histopathological abnormalities and cervical cancers between the studies and Medicare beneficiaries?	We did not find direct evidence on the spectrum and prevalence of histopathological abnormalities and cervical cancers in Medicare beneficiaries. We also could not answer this question reliably directly from the primary studies we reviewed.  We have changed the text accordingly and removed the prior text as it did not directly answer the question.
1	Executive Summary	The conclusions seem justified given the data presented.	Thank you.
1	Executive Summary	I will not repeat comments in the following sections, if they already came up in the Executive Summary.	Acknowledged. We have edited the main report accordingly.
2	Executive Summary	The Executive Summary is clear and complete in its report of the background issues, the methods of the review, the results and conclusions reached.	Thank you.
3	Executive Summary	[Referring to comment about page 13- next comment in table] See comments for suggested re-wording. Histology is not needed for the accepted I option of a “see and treat” approach, and histologic	We have incorporated the “see and treat” concept.

	ES-2; Line 9	confirmation has a subtly different utility (allowing lesions to spontaneously regress and missing small lesions that apparently are not clinically significant) This is important because a definitive diagnosis of HSIL IS a definitive diagnosis with only about 3% interpretive error rate. That is to say, cytologic diagnoses of HSIL without histologic confirmation are rarely false-positive results.	
3	Executive Summary	<p>The text on page 13 (of the document, not the page number) reads: “Generally, a higher grade of cytology indicates a greater risk for higher classes on subsequent histology, but abnormal cytology may also be associated with both more or less severe histologic findings. <b><i>This is why histology is needed for definitive diagnosis</i></b>” (emphasis is mine).</p> <p>This paragraph is not accurate and in my opinion should be changed to the following:</p> <p>Generally, a higher grade of cytology indicates a greater risk for higher classes on subsequent histology, but abnormal cytology may also be associated with both more or less severe histologic findings. <b><i>Historically, all treatment decisions are based on histology: this approach allows equivocal cytologic findings to be confidently diagnosed as a high grade lesion requiring ablative therapy (LEEP, conization or Cryosurgery). Histologic confirmation of a definitive high grade lesion on cytology is also useful by decreasing the number of unnecessary ablative procedures since some high grade lesions will regress spontaneously, or be so small that they cannot be identified on colposcopic sampling and presumably not pose an immediate threat warranting definitive ablative therapies. Approximately 70% of definitive high grade lesions diagnosed by cytology are confirmed histologically. A “see and treat” approach for patients with definitive high grade findings on Pap testing is acceptable in current guidelines. See and treat allows a definitive ablation to take place upon the first colposcopic visit. This approach is cost-effective by allowing a definitive ablative</i></b></p>	Thank you. Your edits have been incorporated.

		<b><i>therapy without the cost of an additional colposcopic exam and histologic confirmation.</i></b>	
3	Executive Summary  ES-2; Cytologic Screening	[Referring to 3 <sup>rd</sup> paragraph of section] See above. “See and treat” is an option.	“See and treat” was added to this paragraph.
3	Executive Summary  ES-2; Cytologic Screening	[Referring to 5 <sup>th</sup> paragraph of section] Again, this paragraph does not acknowledge See and Treat approaches.	“See and treat” was added to this paragraph.
3	Executive Summary  ES-3; Current Guidelines for CC	It may be worth pointing out that several clinical trials outside the United States have documented the value of “primary HPV screening”, i.e., if the HPV test is negative, no other testing is needed, and if the test is positive, cytologic Pap testing is effective for triage because of the high specificity of Pap test findings.	This was added.
3	Executive Summary  ES-15 Strength of Evidence  KQ 4	Another complication of an ablative procedure is the potential inability to have an adequate subsequent colposcopy examination. This is significant because anyone who has had an ablative procedure requires prolonged, if not life-long, cervical cancer screening. If a screening test comes back abnormal on a patient who has already had an ablation, it can be difficult to visualize the lesion and allow the lesion to be optimally treated.	This was added to adverse events.
4	Executive	Same comment regarding key questions; initial rationale for why HPV 16 and 18 is used. There are also ISH of p16 INK4. Although this is not a marker of chromosomal damage, neither is HPV 16 or	We narrowed the scope following a horizon scan to focus on probes that had been studied in the

	Summary	18.(unless integrated). There are many other ISH markers, which again are not necessarily markers of chromosomal damage, but the rational for their exclusion needs to be upfront and clearer.	greatest number of studies. (see histogram in appendix)
4	Executive Summary	KQ2g is somewhat confusing since it asks what is the prevalence of the markers detected by ISH by age or race/ethnicity. I don't think the results/discussion at all answers the question since HPV 16 and 18 are in question, not general HR HPV.	We conducted a focused search for literature on the prevalence of the markers detected by ISH (i.e., TERC, MYC, HPV16, 18) by age or race/ethnicity in the United States. We did not find direct evidence.  We have changed the text accordingly and removed the prior text as it did not directly answer the question.
4	Executive Summary	I think sensitivities and specificities should be included in the section labeled: FISH for TERC or MYC versus other tests.	There is only one study providing data for each comparison of test combinations in this section, which are shown in the plots of sensitivities and specificities. Thus we did not also include them in the narrative.
4	Executive Summary	The conclusion should include some rational as to why FISH and ISH were even investigated. As I mentioned, the lack of validity of these in the scientific community is already well known. What was the possible rational to investigate FISH/ISH.	Thank you.  We added this sentence to the conclusion:" Although ISH tests are marketed by some laboratories for triaging women with abnormal screening tests, there is a lack of standardization of probes and procedures that needs to be addressed. "
5	Executive Summary	Generally, the term HPV (geno)types is more common than "strains"	The term has been changed according to your suggestion.
5	Executive Summary	In the ES and throughout the whole document, the use of cytology terminology is not clear. In the most recent Bethesda nomenclature (that is referred to in the report), ASCUS is separated into ASC-US and ASC-H. Previous reports from older studies, e.g. ALTS, use ASCUS. The authors need to specify which group they	We added in the methods for KQ3a: While the Bethesda nomenclature has evolved to divide ASCUS further into ASCUS and ASC-H, we were limited to using ASCUS as used in the studies. If the study

		refer to.	differentiated ASC-US and ASC-H, we included results for ASC-US only.
5	Executive Summary	There has been a recent update in nomenclature of cervical histology (LAST effort, Darragh et al. 2012). This new nomenclature should be referred to when discussing histology endpoints.	We added in Methods for KQ3a that: Studies included in this review had not uniformly adopted the recently published LAST nomenclature to categorize histology into high grade CIN or low grade CIN.
5	Executive Summary	Integration is not necessary for progression to invasive cancer. Many cancers harbor episomal HPV only (Vinokurova et al. 2009).	This sentence was edited accordingly.
5	Executive Summary	The reference for HPV genotype prevalence in cancers should be updated to the most recent IARC report (Li IJC).	We provided a reference by IARC (Clifford 2006)
5	Executive Summary	A cytology result of HSIL is sometimes used to make treatment decisions without biopsy results. Thus, the statement that cytology alone cannot be used for treatment decisions is incorrect.	This section was revised and this sentence was deleted. We added that, with HSIL on cytology, a see-and-treat approach may be taken.
5	Executive Summary	The description of the recent ACS screening guidelines update should clearly state that co-testing with 5 year screening intervals was recommended as preferred option over cytology at 3 year intervals.  In general, the description of the recent screening guidelines update should be more precise throughout the document.	This was done.  The description was revised.
5	Executive Summary	Data sources: It is not clear whether abstracts from the HPV conferences and Eurogin meetings were included in the data sources, as these meetings would be most likely to include this topic.	We reviewed abstracts from the past two years from ACOG and the American Society of Clinical Oncology. In addition we reviewed the past year of the ASCCP. This is stated in the Methods.
6	Executive Summary	A few clarifications regarding the absolute nature of how laboratories classify and report cervical lesions. Most (but not all) laboratories use the Bethesda terminology. It's not an absolute requirement, but a recommendation. The table demonstrating the Bethesda classification is not complete- would be best to indicate	We added that most laboratories use the Bethesda System.  We now list the complete classification.

		as such if the entire classification is not displayed.	
6	Executive Summary	New recommendations (CAP-ASCCP) for reporting terminology for HPV- related lower genital tract lesions have just been released which include cervical biopsies. The new terminology classifies in a two-tiered system (LSIL and HSIL- similar to PAP testing classification). The changes are not broadly used yet (and probably not broadly accepted) but exists as the new recommended reporting terminology. The CIN definition is important because the pathologist must still determine what proportion of the epithelium is involved (lower 1/3, lower 2/3, or full thickness) and then stratify to LSIL or HSIL. The CIN2 lesions would be stratified based on p16 immunohistochemical results.	Thank you. This was incorporated.
6	Executive Summary	In the executive summary- it's more important to define the CIN terminology by proportion of epithelial involvement rather than "degree of atypia". It's more accurately described in the expanded background section.	We included the full description from the expanded background section.
6	Executive Summary	In the fourth paragraph under "Cytologic Screening for Cervical Cancer" the sentence starting with "HPV testing detects....." is misleading. Stating that HPV testing detects HPV DNA before cellular abnormalities are evident implies that all HPV infections follow with the development of cellular abnormalities. It also understates the fact that some cervical lesions are detected by PAP testing alone- even when HPV testing is negative. Better to just simply state that HPV testing detects the presence of HPV DNA, even though cervical abnormalities may not be present or may not develop.	Sentence has been moved to the HPV section, and your suggestion for editing has been incorporated.
6	Executive Summary	Under principles of ISH- the first sentence states that the chemical tag is detected by the "technician". Many methods of detection exist- including some that are automated visual detection systems. Any ISH final interpretation must be rendered by a pathologist (not just a technician). Would be better to state that the chemical tag is detectable by "various methods" and interpreted by a pathologist.	The section has been revised and this sentence was added: "Any ISH final interpretation must be rendered by a pathologist."
1	Introduction/ Background	Usually, HPV strains are called "types".	This has been changed consistently to genotypes.
1	Introduction/	There is very little evidence that Chlamydia is a risk factor for HPV persistence. HSV is no longer thought to be a co-factor at all.	The reference to chlamydia was removed.

	Background		
1	Introduction/ Background	Other co-factors for cancer given infection are multiparity and long-term OC use, and so far there has not been data proving at what stages in the carcinogenic process the co-factors act.	This was added: Other co-factors for cancer given infection are multiparity and long-term oral contraceptive use, but so far it is unclear in what stages in the carcinogenic process the co-factor act.
1	Introduction/ Background	The statement is made that cervical cancer mortality peaks in middle ages, then rates are quoted that show increases in later age groups. This seems inconsistent.	We deleted this section.
1	Introduction/ Background	The high vs. low risk HPV table is not consistent with IARC classifications. I disagree with some of the attributions.	Our list has been revised.
1	Introduction/ Background	No one knows the prevalence of HPV of the beta and gamma strains, combined with alpha, in the US population. There is no valid way to generate this estimate, currently.	We agree, thank you.(pertains to question 2g)
1	Introduction/ Background	The treatment of CIN3 lists radical hysterectomy, chemotherapy, etc. These are treatments for cancer, not CIN3.	The sentence has been edited accordingly.
1	Introduction/ Background	The physical risks of colposcopy are not correctly stated, do the authors mean the risks of treatments?	The sentence has been edited accordingly.
1	Introduction/ Background	ACOG participated in the ACS/ASCCP/ASCP meeting, which came after the 2009 ACOG guidelines and might alter them in the future.	We have removed this reference to the outdated guidelines.
1	Introduction/ Background	The guidelines for HPV negative ASC-US are actually different than said here (this combination is treated like a dual negative co-test). This topic is taken up again in the 2012 update meeting just concluded in Bethesda by ASCCP and other organizations.	We changed this based on the November 2012 ACOG guidelines: Women with ASCUS on Pap testing and a negative HPV test continue with routine screening as indicated for their age.
2	Introduction/ Background	The background presented in the report is accurate and consistent with our current knowledge of cervical carcinogenesis, management protocols that are currently in place, and the assumptions of the methods by which ISH would be used in practice.	Thank you.

3	Introduction/ Background- Page 2; 2 <sup>nd</sup> p	See previous note. This section of the manuscript should be incorporated into the previous section with the edits suggested. References also need to be merged and updated.	The 34 <sup>th</sup> page of the document marks the beginning of the main report. The prior 34 pages contain the abstract and executive summary. Executive summaries provide a concise summary of the report. All edits have been incorporated throughout.
3	Introduction/ Background- Page 5; 2 <sup>nd</sup> p	See previous note. This section should be incorporated into the previous section and references updated.	The 34 <sup>th</sup> page of the document marks the beginning of the main report. The prior 34 pages contain the abstract and executive summary. Executive summaries provide a concise summary of the report. All edits have been incorporated throughout.
3	Introduction/ Background- Page 6; Last P	See previous notes in this redundant section. Risk of inadequate subsequent colposcopy...	The 34 <sup>th</sup> page of the document marks the beginning of the main report. The prior 34 pages contain the abstract and executive summary. Executive summaries provide a concise summary of the report. All edits have been incorporated throughout.
4	Introduction/ Background	<p>I think the section on FISH/ISH should be expanded to include other biomarkers available/tested, (e.g. p16, etc) and the limitations of ISH (reliability, staining defects, etc).</p> <p>Difficulty in distinguishing CIN 2 from CIN 3 and CIN 1 should be expanded. Many pathologist do not believe CIN 2 is a lesion and currently use p16 staining to upgrade or down grade.</p> <p>Also, CIN 2 by many is not considered a good clinical endpoint, only CIN 3. These points should be expanded to better understand how biomarkers staining remains critical to more efficient pathology diagnosis.</p>	<p>The section on ISH testing was edited to address reliability.</p> <p>The nomenclature by LAST and the recommendation to use p16 to triage CIN2 to either high grade or low grade CIN has been added to the background. P16 staining uses immunohistochemistry and was therefore not included in the section on ISH testing.</p> <p>While CIN3+ is a more meaningful endpoint, not all studies reviewed provided analyses for CIN2+ and for CIN3+ separately. Te LAST recommendations endorses use of p16 staining to clarify the grade for</p>

			<p>CIN2, however this was not an established approach in the studies reviewed which generally predate the new nomenclature.</p> <p>We have added to future research recommendations that future studies need to be bigger to yield more precise estimates for HSIL.</p>
5	Introduction/ Background	Several points mentioned above are repeated here and apply equally (e.g. mention LAST, role of integration, cytology for therapy decisions)	We addressed these issues consistently throughout.
5	Introduction/ Background	The most recent IARC classification is different from what is shown in table 3.	Table 3 has been revised.
5	Introduction/ Background	Page 6, last paragraph- the complications described are not related to colposcopy, but to treatment!	This sentence has been edited accordingly.
5	Introduction/ Background	Page 9, third paragraph: CIN3 is considered the best surrogate of cancer risk, a substantial proportion of CIN2 regresses spontaneously	The sentence has been edited to include only CIN3.
6	Introduction/ Background	The tables showing Bethesda classification and Cervical Histology classification: It's best to state that the Bethesda classification shown is only a portion of the entire classification system. The Cervical Histology classification in the table is not one that is used. "Borderline" is not used as a specific classification, nor is "Low- or high-grade cervical glandular intraepithelial neoplasia (CGIN). This may be published but certainly not a system that is accepted for broad use. "Adenoglandular" carcinoma is not a term used in current pathology diagnoses.	We are now showing the full Bethesda classification for cytology and the LAST recommendations for histology.
6	Introduction/ Background	The first paragraph under natural history of Cervical Cancer- the anatomy of the cervix with squamous and glandular cells is not correctly stated. The cervix is the lower narrow portion of the uterus that consists of an extocervix, lined by squamous epithelium and an endocervix, lined by glandular epithelium. The ectocervix transitions to the vagina inferiorly. The endocervix is superior to the ectocervix and transitions to the endometrial cavity of the uterus. The transition from ectocervical lining to endocervical lining is called the transformation zone.	Sentence has been edited accordingly.

6	Introduction/ Background	Under the epidemiology of Cervical Cancer: adenocarcinoma and the precursors are harder to detect than squamous carcinoma and precursors, because distinguishing features for glandular lesions are more subtle and challenging to interpret. Limited sampling may also play a role.	Sentence has been edited accordingly. These points were incorporated.
1	Methods	I am struck by the large number of exclusions of articles. As one focused question, why was it necessary to have a restriction to ASC-US and LSIL cytology? As mentioned, another possible application, where ISH is an alternative to HPV16/18 genotyping using non ISH methods, is for the HPV-positive cytology-negative woman. Why did the review focus narrowly when the data are already so sparse? I would have expected some treatment of the broader literature on ISH and cervical specimens.	The question was focused on the most clinically meaningful clinical scenarios. This is what FISH tests are currently marketed for.  We added that no studies provided results for ISH in the HPV-positive cytology-negative woman.
2	Method	The methods used are consistent with those used in other reviews of this type. The literature search, exclusion/inclusion criteria, strength assessment and grading of bias were all consistent with current reviews of this type and appear appropriate to the task at hand.	Thank you.
4	Method	relatively straightforward	Thank you.
5	Method	A stronger focus on evaluation of ISH slides would be desirable.	Key Questions 2b–f, focused on narrative or quantitative information on reliability and reproducibility of ISH tests and possible factors interfering with analytic test performance. However, we found only sparse data in the reviewed studies on pre-analytic issues and how they impact analytic validity.
6	Methods	No comments.	Thank you.
1	Results	The handling of the data that were used seemed correct to me.  The one concern I had was the generation of summary estimates for sparse data with major heterogeneity. It is nice to have summary estimates, but when the study variation is this extreme, it is hard to trust them.	Thank you.  As above, we agree that the data are statistically heterogeneous. This heterogeneity is taken into account by the random effects model we used (resulting in “appropriately” wide confidence

			intervals for the summary estimates of sensitivity and specificity). However, we do not base the decision to perform meta-analysis on observed statistical heterogeneity. The studies we synthesized were conducted in women with ASCUS or LSIL, using FISH as the index test with a histological diagnosis of CIN2+ or CIN3+ on biopsy. On the basis of these similarities we judged that the studies could be quantitatively synthesized.
2	Results	The results are presented in a clear and concise fashion and are again consistent with my own current assessment of the use of ISH in cervical cytology (from a similar body of literature).	Thank you.
4	Results	see below table 4: Table 4; it is unclear from this table if the probes for HPV 16 and 18 were reported separately or as a group (high risk vs low risk)."	We inserted the following clarification: ISH assays tested for variable combinations of HPV genotypes; some tested for a panel of high risk HPV, and some tested for specific genotypes. Studies did not clearly identify the specific origin of a positive test.
5	Results	Several times in the document, it is mentioned that HPV PCR detects mRNA rather than DNA. This is incorrect, most HPV PCR assays detect viral DNA.	The document was edited as follows:  "Staining on ISH tests identifies episomal or nuclear HPV DNA. Most HPV PCR assays detect viral DNA, but PCR does not differentiate between episomal or nuclear DNA."
5	Results	Beta-glucan is mentioned as a sample adequacy control- I assume you mean beta-globin	Glucan was changed to globin.
6	Results	No comments	
1	Discussion/ Conclusion	Based on the results, the skeptical assessment seems correct. Without sufficient reliability data, and with inadequate assessment of analytic and clinical performance, there is no other reasonable conclusion; the techniques cannot be recommended for clinical use. The scant data do not even seem that promising for the	Thank you.

		applications, in that the sensitivity/specificity trade-offs are not that good (if one accepts the pooled estimates). Very high sensitivity for CIN3+ is required for triage, higher than the estimates presented for the chromosomal markers. For HPV16 and HPV18, in light of the data I do not see an advantage to ISH compared to other well validated methods.	
2	Discussion/ Conclusion	The conclusions of the review, that ISH is not ready for routine use in the triage of cervical cytology abnormalities (ASC-US and LSIL), appear to be well-justified from the literature identified in the review. The low number of studies, the low number of specimens in each study, and the highly variable protocols do not lend confidence that the accuracy of either a positive or a negative test can be relied upon beyond what has already been shown to be an effective triage method. Also, as is noted in the discussion, the playing field for triage is changing rapidly. HPV genotyping creates a different landscape for case triage and it is not clear whether ISH for any of the studied markers will provide better triage than genotyping alone. Any proposed application will need to be studied in the context of these new developments. Further study of these methods in well-designed, well-powered, and unbiased studies will be required to decide about their ultimate utility.	Thank you.
4	Discussion/ Conclusion	Agree with conclusion that there is no evidence that these assays should be used clinically.	Thank you.
5	Discussion/ Conclusion	The discussion should also address evaluation of ISH slides.	We have expanded the research gaps section to comment on the need to establish common measures and standards for ISH tests.
6	Discussion/ Conclusion	Agree with discussion and conclusions	Thank you.
2	Tables	The tables are clear and relevant.	Thank you.
4	Tables	Table 4; it is unclear from this table if the probes for HPV 16 and 18 were reported separately or as a group (high risk vs low risk).	We inserted the following clarification: ISH assays tested for variable combinations of HPV genotypes; some tested for a panel of high risk HPV, and some

			tested for specific genotypes. Studies did not clearly identify the specific origin of a positive test.
6	Tables	No comments	
2	Figures	The figures are clear and relevant.	Thank you.
3	Figures	Figure 4, from page 47 (see below, middle branch) recommends Pap testing after HPV 16/18 positivity, but this could not make sense. Regardless of the Pap test result, all patients should go to colposcopy. Also, in the upper branch of Figure 4, testing for HPV 16/18 following a high risk positive screen and an LSIL Pap test result is not predicted to have value, based on current evidence. Such patients have roughly a 20% chance of a CIN 2+ and need colposcopy regardless of HPV 16 or 18 positivity.	[NOTE: this comment refers to the 47 <sup>th</sup> page of the document, not the page labeled as 47.]  As indicated in the title of the figure, these are hypothetical scenarios for primary screening for HPV that do not depict current standard of care. However, we believe that these branches may be hypothetical options. There is a precedent for older women with LSIL to be tested for HPV 16/18.
4	Figures	Good	
5	Figures	Figure 4: Primary HPV testing is not recommended for women 21 years and older, but for women 30 years and older.	The title of Figure 4 was edited to indicate this.
5	Figures	Figure C1- it seems that the chr.7 probe was analyzed more frequently than MYC- why was it not included in the review?	The chromosome 7 probe was used as a control and not a marker of damage so while we included studies that looked at this probe as part of a panel, we did not believe it should be the focus of this review.
6	Figures	No comments	
1	Appendices	HPV53 and HPV66 are no longer considered HR, based on improved data from IARC and ICO.	HPV53 is now listed in with the low-risk group. The IARC report still lists HPV66 as high risk.
2	Appendices	No further comment.	

4	Appendices	no comment	
5	Appendices	The heading “FISH Results Not Reported for Cytology Classification According to Histology Classification” is unclear- some of the articles seem to provide information about the diagnostic use of ISH assays, e.g. Wright 2012	This title was edited accordingly.
1	References	No comment, except that the general review references are limited explaining some of the errors in those general sections.	This section was updated.
2	References	No further comment – I am not aware of any significant study not captured by the review.	Thank you.
6	References	No comments	
4	Appendices	no comment	
6	Appendices	No comments	

<sup>1</sup> Peer reviewers are not listed in alphabetical order.

<sup>2</sup> If listed, page number, line number, or section refers to the draft report.

<sup>3</sup> If listed, page number, line number, or section refers to the final report.

Project Name: **In-Situ Fluorescent Hybridization (FISH) or Other In Situ Hybridization (ISH) Testing for Chromosomal Damage in Uterine Cervical Cells to Predict Potential for Dysplasia/Malignancy**

Project ID: **CANC0511**

Table 2: Public Review Comments

Reviewer Name <sup>1</sup>	Reviewer Affiliation <sup>2</sup>	Section <sup>3</sup>	Reviewer Comments	Author Response <sup>4</sup>
Thomas Ried	National Institutes of Health	General	<p>It is with great interest that we have read your Technology Assessment Report on the use of ?Fluorescence In Situ Hybridization (FISH) or Other In Situ Hybridization (ISH) Testing for Chromosomal Damage in Uterine Cervical Cells to Predict Potential for Dysplasia/Malignancy?, currently posted for public review. Such novel method evaluations are important and helpful to advance medical practice, and we are impressed by the comprehensive and objective way in which this report was compiled. While we largely concur with the authors? conclusions as to the use of FISH protocols for HPV detection, we would like to provide comment regarding the use of visualization of chromosome 3q gain, and with it amplification of the human TERC gene for the detection of cervical cancer and the assessment of individual progression risk of cervical dysplastic lesions to invasive disease.</p> <p>The utilization of a FISH test for the detection of extra copy numbers of the 3q chromosome arm was initially triggered by a publication from the National Human Genome Research Institute/NIH/DHHS in 1996. At that time, this observation was considered to be the missing link in cervical carcinogenesis. Subsequent to that discovery, which was patented, several groups, including the NHGRI team, proceeded to develop a diagnostic test for visualization of 3q gain directly in routinely collected cytological samples (pap smears). The application, as documented in multiple publications, unambiguously confirmed that the gain of the chromosome 3q arm predicts the progression of premalignant lesions to invasive disease. The power of this genetic marker was moreover replicated in multiple independent studies, not all of which were taken into consideration in the TA (PMID:15793301, 14507648).</p>	Thank you.

		<p>We agree with the notion that at the present time a FISH test cannot serve as a practical initial population based screening tool for cervical cancer. However, we submit that based on substantial public evidence, its value for triage (i.e. for discerning low risk or high risk for progression to invasive disease) cannot be questioned. This assessment is based on i) a retrospective study (PMID:15793301) and ii) on a prospective study (PMID: 19880826) that compared FISH results on pap smears with the results obtained from the histological evaluation of follow-up cervical biopsies.</p> <p>The identification of test standards and cutoff definitions is arguably a challenge for any diagnostic test, as exemplified in the ongoing debates surrounding e.g. serum PSA measurement in prostate cancer, or the determination of HER2 amplification in breast cancer.</p> <p>The reason that there is apparently limited strength of evidence in essentially all of the questions in your analysis is precisely due to the fact that there are no commonly agreed upon test and validation standards yet. Nevertheless we feel that there are now a considerable number of studies that provide very compelling data for further rigorous assessment and broader discussion on testing guidelines. These difficulties, however, do not disqualify the value of a test. <b>A constructive solution would be to convene a conference with experts involved in the field, to agree on common measurement guidelines, a path that was successfully pursued to arrive at the consensus Bethesda classification for cytological abnormalities in cervical cancer.</b> We do sincerely believe that the implementation of a molecular marker for progression risk stratification would not only improve prognostication but would also help reduce over-treating women with unnecessary invasive procedures. The question cannot be whether 3q testing should be implemented but how it will best be standardized for clinical use.</p>	<p>We added a sentence in the discussion/research gaps section in response to the comment in bold and highlighted suggesting an expert conference to agree on common measurement guidelines. This sentence reads: "An expert conference may be helpful to agree on common measurement guidelines, a path that was successfully pursued to arrive at the consensus Bethesda classification for cytological abnormalities in cervical cancer."</p>
--	--	--	--

<sup>1</sup> Names are alphabetized by last name. Those who did not disclose name are labeled "Anonymous Reviewer 1," "Anonymous Reviewer 2," etc.

<sup>2</sup> Affiliation is labeled "NA" for those who did not disclose affiliation.

<sup>3</sup> If listed, page number, line number, or section refers to the draft report.

<sup>4</sup> If listed, page number, line number, or section refers to the final report.