Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: ***Emerging Approaches to Diagnosis and Treatment of Non-Muscle-Invasive Bladder Cancer***

Draft review available for public comment from November 20, 2014 to December 10, 2014.

Research Review Citation: Chou R, Buckley D, Fu R, Gore J, Gustafson K, Griffin J, Grusing S, Selph S. Emerging Approaches to Diagnosis and Treatment of Non-Muscle-Invasive Bladder Cancer. Comparative Effectiveness Review No. 153. (Prepared by the Pacific Northwest Evidence-based Practice Center.) AHRQ Publication No. 15(16)-EHC017-EF. Rockville, MD: Agency for Healthcare Research and Quality. October 2015. www.effectivehealthcare.ahrq.gov

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site in draft form for public comment for a 3-week period. Comments can be submitted via the EHC Program Web site, mail or E-mail. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft 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.

Source: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2137
Published Online: October 27, 2015

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #1	General Comments	This document does not go into the difficult controversy regarding cost of urinary biomarkers, fluorescence, and the cost effectiveness of varying intervals of cystoscopy.	Evaluations of cost and cost-effectiveness were outside the scope of this review. However, we revised the Discussion/Implications for Clinical and Policy Decisionmaking section to note that cost may impact decisions regarding use of biomarkers and fluorescent cystoscopy.
TEP Reviewer #1	General Comments	Can specific questions re: markers be answered or are they answerable. Important queries include: --how sensitive/accurate are urinary biomarkers in detecting urothelial cancer in the previously undiagnosed patient/screening patient. --how often and how long should the interval use of cystoscopy be performed based on either pre-test probabilities or patient characteristics. --how often and how long should cytology be performed in evaluation and followup of screening patient and patient with known bladder cancer and can we differentiate based upon clinical features such as grade and stage or primary vs recurrent or receiving or not receiving chemotherapy.	As noted in the methods, we excluded studies of biomarkers for screening (i.e. for diagnosis of bladder cancer in patients without symptoms or prior bladder cancer). Results were stratified for patients with symptoms and for surveillance. The frequency and duration of cystoscopy and cytology were not Key Questions for this review. However, we are not aware of any studies evaluating effects of differences in cystoscopy or cytology intervals of duration of testing on clinical outcomes. To clarify, this report does not provide clinical

Commentator & Affiliation	Section	Comment	Response
			recommendations, rather its purpose is to synthesize evidence.
TEP Reviewer #1	Methods	Reasonable in light of funding constraints	Thank you for the comment.
TEP Reviewer #1	Results	Should examine impact of EMDA (electromotive therapy) and Synergo (TM) microwave therapy when used in conjunction with intravesical treatments. Data exists BUT studies done in Europe.	These technologies were excluded because they are not FDA approved and are not in common use in the United States, according to the TEP.
TEP Reviewer #1	Results	For mitomycin C specifically, the data should be separated for effectiveness as well as for side effects--the impact of perioperative single dose vs induction course/maintenance course. There is in the table the comparison trial but I would want to know versus no treatment as well. I don't seem to find data showing effective of single mitomycin c as I know there are a couple	Results for MMC vs. TURBT alone are presented in KQ 2 (p 29-30). The draft noted that estimates were similar in subgroup and sensitivity analyses; we revised to present the estimates for single instillation vs. induction/maintenance therapy. Side effects of MMC were very limited from a single trial and were presented in KQ 8 (p 87-88).
TEP Reviewer #1	Results	On page 40, was the comparison of MMC vs 1-2 weeks later--single dose vs single dose or single dose vs a weekly course of medication?	It was 9 instillations over 6 months on the day of TURBT vs. starting 1-2 weeks after TURBT (see p 74). We revised the SOE table to make this clearer.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #1	Discussion/ Conclusion	Would mention the importance of improved/more uniform complications reporting.	Revised as suggested (p 221 line 18).
TEP Reviewer #1	Clarity and Usability	can the data regarding risk stratification be evaluated and included. this would include the EORTC risk calculator and CUETO risk calculators -- how do these classification systems integrate with treatment modalities?	KQ 2 addresses evidence on effects of using a formal risk-adapted approach on clinical approach (no studies). KQ 4b and 4c address how comparative effectiveness varies according to tumor and patient characteristics. The report does not make clinical recommendations.
TEP Reviewer #1	Clarity and Usability	how do the authors believe a blinded fluorescence trial be accomplished today and get IRB approval when the treatment is standard of care in the US and randomized trials (with possible bias as described) have shown a benefit compared to white light alone? Would such a trial accrue?	As noted in KQ 6, a trial was performed in which fluorescent cystoscopy was used in all patients, but the cystoscopist was blinded to whether a photosensitizer or placebo was instilled. We believe similar trials can be performed and given the negative results of the above trial are appropriate to further delineate the role of fluorescent cystoscopy.
TEP Reviewer #2	General Comments	This report appears to thoroughly discuss the diagnosis and treatment of NMIBC. The conclusions validate or at least are consistent with general practice.	KQ 4 addressed the comparison of radiotherapy vs. intravesical therapy or radiotherapy vs. cystectomy. The only trial compared

Commentator & Affiliation	Section	Comment	Response
		For the radiotherapy key question 4, the issue seems a bit off target. The key question looks at the role of radiotherapy versus intravesical therapy, however, the actual key question is: for patients who have failed intravesical therapy and still have NMIBC, is radiotherapy a valid alternative to radical cystectomy? However, the report overall seems to step back from discussing the management of patients who have failed intravesical therapies and are confronting the option of radical cystectomy. The EAU guidelines referenced in your report (Eur Urol 64:639, 2013) note in Table 8 options for BCG failure and recurrences after BCG including bladder preserving strategies, i.e. radiotherapy, for patients not suitable for cystectomy.	radiotherapy vs. no radiotherapy (for lower risk lesions) and radiotherapy vs. intravesical therapy. We revised KQ 3c to include a discussion of two trials of different intravesical therapy regimens in patients with recurrence after BCG. We did not identify any trials comparing bladder preserving strategies versus radical cystectomy in patients who failed BCG (KQ 4); we added this as a research gap in the Discussion.
TEP Reviewer #2	Discussion/ Conclusion	The research section is clear, however, it is unclear why the quality of research performed to date is so poor. Why don't we have a better data repository in order to answer the key questions?!	Thank you for the comment.
TEP Reviewer #2	Discussion/ Conclusion	Additionally, as mentioned above, further research into the BCG recurrent/resistant patient would be valuable as this is a clinically difficult situation. Cystectomy now? Cystectomy later? Alternatives to cystectomy?	We revised KQ 3c to include a discussion of two trials of different intravesical therapy regimens in patients with recurrence after BCG. We did not identify any trials comparing bladder

Commentator & Affiliation	Section	Comment	Response
			preserving strategies versus radical cystectomy in patients who failed BCG (KQ 4); we added this as a research gap in the Discussion.
TEP Reviewer #2	Clarity and Usability	Yes. Well structured and easy to follow.	Thank you for the comment.
TEP Reviewer #3	General Comments	The report is a very good summary of current literature with regards to clinically available markers and intravesical agents. The key questions are appropriate. There are 2 areas that were not specifically discussed. 1. The use of immediate post-TURBT intravesical therapy. 2. Options for treatment after failure of initial intravesical therapy such a BCG. The latter is an area of significant clinical need and confusion with regard to best treatment. The cost implications of added technology and markers was not discussed and considering the fact that bladder cancer is one of the most expensive diseases (as noted in intro), this is a factor for health care decisions.	We revised the Results to more clearly present the results of trials of single instillation therapy. We also revised the Results (KQ 3c) to highlight results of trials of patients that failed BCG. Evaluations of cost and cost-effectiveness are beyond the scope of this review, though we revised the Discussion/Implications for Clinical and Policy Decisionmaking section to note that cost may impact decisions regarding use of biomarkers and fluorescent cystoscopy.
TEP Reviewer #3	Introduction	Well written	Thank you for the comment.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #3	Methods	the criteria used as justifiable and outcomes are appropriate. The statistical methods are appropriate.	Thank you for the comment.
TEP Reviewer #3	Results	There are several areas that could be improved. Some of the studies such as ref 116 for gemcitabine is one of a single instillation trial. One needs to go to pg 85 of the report after having read 2 summaries ES 11 and pg 28 to realize that the conclusion that gemcitabine is not more effective than no therapy is based on a single dose. This may not be true for multiple doses as SWOG trial S0353 (J Urol. 2013 Oct;190(4):1200-4. doi: 10.1016/j.juro.2013.04.031. Epub 2013 Apr 15.) showed activity in BCG failure patients when given 6 doses with maintenance. It is not easy to know if single postop dose trials are intermixed with full treatment and this can make it confusing for the reader. Ideally a section on immediate postoperative intravesical therapy would be written to separate out these trials.	We revised the Key Point to be clearer that the trial of gemcitabine was a single instillation trial (p 28, fourth bullet). The Skinner study did not meet inclusion criteria because it was an uncontrolled study. We revised the Results for KQ2 to more clearly distinguish results for single instillation versus induction/maintenance therapies.
TEP Reviewer #3	Discussion/ Conclusion	The conclusions are appropriate considering the heterogeneity of the data and studies that were available. There are significant limitations to any analysis that can be performed especially with markers. There might be a role for discussing verification bias regarding the marker	We revised the Discussion (Limitations of the Evidence Base section) to include a discussion of the potential impact of verification bias. We revised the Results and Discussion to address the

Commentator & Affiliation	Section	Comment	Response
		studies since patients with positive markers and normal cystoscopy never underwent biopsy, the concept of anticipatory positives has emerged with little evidence regarding the significance of a positive marker (true or false positive). This has impacted the ability to interpret and act upon positive markers unlikely cytology. In terms of future research, there could be discussion of treatment options in patients who recur despite intravesical therapy. This is considered an area of significant need but is not addressed. There is also room for assessing utility of markers in certain clinical situations such as patients with abnormal cystoscopy or atypical cytology for which there are several studies especially with FISH.	area of treatment in patients who recur despite intravesical therapy. We also revised the Discussion to mention potential uses of biomarkers in patients with abnormal cystoscopy or atypical cytology.
TEP Reviewer #3	Clarity and Usability	I think the report highlights more of the limitations of the literature than gives guidance on usage of markers or fluorescent cystoscopy. It will possibly inform policy but in some ways it may limit use of newer markers and technology because of the uncertain benefits of existing tools (markers/blue light,etc)	Thank you for the comment. The purpose of the report is not to provide clinical recommendations, but to synthesize the evidence.
TEP Reviewer #4	General Comments	Yes	Noted.
TEP Reviewer #4	Introduction	Well written and comprehensive	Thank you for the comment.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #4	Methods	yes	Noted.
TEP Reviewer #4	Results	yes	Noted
TEP Reviewer #4	Discussion/ Conclusion	Yes	Noted.
TEP Reviewer #4	Clarity and Usability	Yes	Noted.
TEP Reviewer #5	General Comments	Overall, I thought the review was very well done and an excellent review of the literature.	Thank you for the comment.
TEP Reviewer #5	General Comments	I was concerned that the results repeated multiple times, which lessened the readability of the document. However, the tables and figures were very useful summaries of the information. The authors also did a good job with the final summaries of the data. These summaries help with final interpretation of the results.	Thank you for the comment. We followed standard AHRQ report formats in terms of presenting results etc in the Executive Summary, Strength of Evidence Tables, Results, and Key points.
Peer Reviewer #1	General Comments	The report is comprehensive, relevant and well done. The Key questions are clear and appropriately answered to the best level possible.	Thank you for the comment.
Peer Reviewer #1	Results	yes, it is well done.	Thank you for the comment.
Peer Reviewer #1	Discussion/ Conclusion	Throughout the manuscript the authors refer to BCG as being associated with chemical cystitis. It is a relatively minor point but BCG causes a granulomatous cystitis. Chemical cystitis is a	We revised to the use the term “granulomatous” cystitis in reference to BCG. We did not use the term “cystitis” alone since it could

Commentator & Affiliation	Section	Comment	Response
		chemotherapy side effect. It would be better to just use the term "cystitis" as a side effect. This problem exists throughout the manuscript.	be confused with infectious cystitis.
Peer Reviewer #1	Clarity and Usability	yes. The report is redundant between the executive summary, body of the manuscript etc. This makes it very repetitive .	The Executive Summary is meant to be a standalone document, so redundancy with the full report is necessary.
Peer Reviewer #2	General Comments	This systematic review provides quite valuable current information for a very important public health issue, i.e. nonmuscle invasive bladder cancer. I believe that the most timely and useful information concerns the analysis of the test performance characteristics of urinary biomarkers and the utility of enhanced cystoscopy. This is because currently existing guidelines (AUA, EUA, NCCN) do not sufficiently address these emerging technologies. This data regarding biomarkers and enhanced cystoscopy will be immediately valuable to clinicians, patients, payers, and policy makers. The systematic review concerning intravesical therapy however, adds little to existing guidelines and is presented in a manner that is in my opinion less useful. This is probably largely a function of the deficient literature regarding this topic which is a recurrent theme in existing systematic reviews.	Thank you for the comment. For intravesical therapies, we addressed the key questions on comparative effectiveness and harms based on the available evidence.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Introduction	very well done; no issues	Thank you for the comment.
Peer Reviewer #2	Methods	The methodology in all respects is state of the art for outcomes research.	Thank you for the comment.
Peer Reviewer #2	Results	I believe that the results of the 3 areas of review (urinary biomarkers, enhanced cystoscopy, and intravesical therapy efficacy) are well presented with adequate description of the studies. In additon, the key findings are clear and apparent. Especially helpful are the wealth of figures and tables.	Thank you for the comment.
Peer Reviewer #2	Discussion/ Conclusion	The discussion and conclusions are well done.	Thank you for the comment.
Peer Reviewer #2	Clarity and Usability	Overall the report is very well done with the major points clearly presented. The conclusions regarding the utility of urinary biomarkers and that of enhanced cystoscopy are especially useful and can be used to help clinicians and policy makers alike. A systematic review of these 2 issues was very much in need.	Thank you for the comment.
Peer Reviewer #2	Clarity and Usability	The analysis and conclusions regarding the efficacy of intravesical chemotherapy as I mentioned above, I find less helpful. The reason for this I think is largely a function of the very deficient literature base in this field. NMIBC is an extremely heterogenous disease. On one end of the spectrum are patients with low grade noninvasive tumors that rarely develop tumors that pose a risk	As described in the Methods and Results (KQ 3), we performed stratified and subgroup analyses of trials based on factors such as tumor stage, grade, presence of CIS, multifocality, and recurrence. In general,

Commentator & Affiliation	Section	Comment	Response
		for progression (and potential mets/death). On the other end are patients with high grade tumors invading the lamina propria with or without coexistant cis who have a life threatening disease from the start. Studies often do not stratify patients according to these tumor characteristics. None the less, practicing clinicians and all the major guidelines approach treatment decisions based on perceived risk of recurrence, progression and potential mortality. The data and conclusions concerning intravesical therapy in this review do not help in this regard.	these factors did not impact findings regarding effectiveness, though estimates were often imprecise. In addition, KQ 3b and 3c address effects of tumor and patient characteristics on estimates of effectiveness, based on within-study subgroup comparisons, again showing little evidence of differences. However, as noted in the Discussion (p 219, lines 32-34) decisions about use of intravesical therapies should take into account the risk of the patient, as benefits are likely to be higher in patients at higher risk, given the potential adverse effects.
Peer Reviewer #2	Clarity and Usability	A further omisson in my opinion, is the lack of attention to data regarding a single perioperative instillation of intravesical chemotherapy (especially for low risk patients).	We revised the Results to more clearly highlight results of single instillation intravesical therapy.
Peer Reviewer #3	General Comments	Report is well prepared and highlights the strengths and weaknesses of the current literature/studies	Thank you for the comment.
Peer Reviewer #3	General Comments	The Key questions are well stated and appropriately answered within the limitations of the data	Thank you for the comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Introduction	Appropriate	Noted.
Peer Reviewer #3	Methods	Appropriate	Noted.
Peer Reviewer #3	Results	Appropriate	Noted.
Peer Reviewer #3	Discussion/ Conclusion	Appropriate	Noted.
Peer Reviewer #3	Clarity and Usability	The report is well structured. The conclusions can be used to inform policy, inform practice decisions and guide future research	Thank you for the comment.
Peer Reviewer #4	General Comments	This is a comprehensive, timely, meticulous and very well done review of the current state of diagnostic and therapeutic tools for patients with non-muscle invasive bladder cancer. As a urologist embedded in this field I provide the following critique and recommendations for consideration of revisions for the final manuscript. I have provided comments specifically for the executive summary and have not repeated them for the full text but request that revisions be made to both.	Thank you for the comment.
Peer Reviewer #4	General Comments	1. Page 12 first paragraph - suggest updating the incidence statistics using 2014 data; relation to smoking should be qualified to reflect cigarette smoking	We revised with updated incidence data and revised to refer to cigarette smoking.
Peer Reviewer #4	Key Question 2	2. Key Question 2 – not sure if this was framed properly as this is what we do each day in practice. We never use a “one size	We revised the KQ to be clearer that it is referring to use of a formal risk adapted

Commentator & Affiliation	Section	Comment	Response
		fits all approach". We do risk adapted therapy with every treatment decision. I think I understand what the question was aiming at but unfortunately one cold conclude that treating every patient with BCG (see below) is appropriate which I believe would send the wrong message.	approach versus not using a formal risk adapted approach.
Peer Reviewer #4		3. Page 26 4th bullet MMC – I disagree with conclusions that SOE is low as this Ph III RCT (Au JNCI 2001) comparing optimized vs. standard showed a clear benefit to the optimized regimen. COI - I was an investigator and co-author of the trial and I am hard pressed to understand why the authors view this well designed and executed RCT does not meet a higher bar.	The risk of bias of the individual study was rated as medium. However, the SOE rates the confidence in the body of evidence and accounts for not only the quality of the individual studies, but also the consistency, directness, and precision. The SOE is rated low because there were only two relatively small trials (n=26 and n=201) with somewhat inconsistent results (one trial reported benefit and the other no benefit).
Peer Reviewer #4		4. Page 28 Question 6 2nd bullet re FC – SOE low – see my comments below regarding my concerns about the SOE.	The Stenzl 2011 trial was the only one in which the cystoscopist did not know whether the patient received intravesical 5-ALA or placebo prior to initial cystoscopy. This design reduces performance bias

Commentator & Affiliation	Section	Comment	Response
			because the cystoscopist does not know whether patients received 5-ALA or not (unlike most trials, in which the cystoscopist is performing white light and fluorescent cystoscopy with 5-ALA, versus white light cystoscopy alone). In the Stenzl 2010 trial, patients were initially randomized to no HAL vs. HAL, and all initially underwent white light cystoscopy. The group randomized to HAL then underwent a second randomization to fluorescent cystoscopy versus no fluorescent cystoscopy. This design does not eliminate performance bias since cystoscopists know what type of cystoscopy they were using during the "second look," though we agree that this design is superior to the standard design, in which patients undergo a single randomization and cystoscopists are unblinded to instillation of the

Commentator & Affiliation	Section	Comment	Response
			photosensitizer. We revised the Results to make this clearer.
Peer Reviewer #4		5. Page 47 line 17-21 –The reference provided (Stenzl et al trial of 5-ALA) was a negative trial and the authors use this to apparently mitigate the body of work showing a reduction in recurrence probability. They base this on the urologist being blinded to the treatment. This is technically correct, as it was a placebo controlled trial but as the same urologist did the procedure under both WL and FC. Therefore they would know if the patients received the ALA based on the presence of fluorescence. The most recent RCT of Hexvix was a positive trial showing both 9 month and long term reduction in recurrence probability (Stenzl, et al 2010 Grossman, et al 2012) providing level I evidence of benefit. On page 48 line 6 the authors state again that the apparent reduction in recurrence probability may be affected by "performance bias". I am not sure I understand what this is. As stated above it is impossible to eliminate potential bias without having two surgeons per case and randomize which one does WL and which does fluorescence. Then the applicability of this to clinical practice is potentially biased, as this two surgeon scenario will never happen. The most	The Stenzl 2011 trial was the only one in which the cystoscopist did not know whether the patient received intravesical 5-ALA or placebo prior to initial cystoscopy. This design reduces performance bias because the cystoscopist does not know whether patients received 5-ALA or not (unlike most trials, in which the cystoscopist is performing white light and fluorescent cystoscopy with 5-ALA, versus white light cystoscopy alone). In the Stenzl 2010 trial, patients were initially randomized to no HAL vs. HAL, and all initially underwent white light cystoscopy. The group randomized to HAL then underwent a second randomization to fluorescent cystoscopy versus no fluorescent cystoscopy. This design does not eliminate performance bias

Commentator & Affiliation	Section	Comment	Response
		recent Hexvix trial, which led to FDA approval controlled for this as best as one can perhaps by a second randomization after WL cystoscopy. This shold also be addressed in the last paragraph page 50.	since cystoscopists know what type of cystoscopy they were using during the "second look," though we agree that this design is superior to the standard design, in which patients undergo a single randomization and cystoscopists are unblinded to instillation of the photosensitizer. We revised the Results to make this clearer.
Peer Reviewer #4		6. Page 47 line 32 – I recommend that a statement be included qualifying the comparison of BCG + maintenance to MMC, noting that no trial or direct comparison has been made to MMC using the optimized regimen described by Au et al JNCI 2001. As noted in the above comment this Phase III trial clearly demonstrates superiority of optimize MMC to a "standard" regimen of 20mg/20cc (1mg/cc) which is the typical comparator in the meta-analyses referenced. See also first bullet page 96 BCG vs. MMC; page 121 line 19	We revised to note that none of the trials of MMC used an "optimized" regimen.
Peer Reviewer #4		7. Page 47 38 line 38 references a review from 2008 and therefore does not include the most recent EORTC trial 30962 comparing one year vs. 3 year	The Oddens trial was included in our review; as described in the Results we believe that the optimum

Commentator & Affiliation	Section	Comment	Response
		maintenance BCG which clearly defines the optimum maintenance schedules in intermediate risk and high risk disease, respectively (Oddens, et al Eur Urol 63:462, 2013; ref 191).	dosing regimen is still uncertain given inconsistent results between trials, including in higher risk patients.
Peer Reviewer #4		8. I am very concerned about the conclusions in the last paragraph that imply that there is limited (or perhaps no role) for intravesical chemotherapy induction with or without maintenance and that BCG is the optimal therapy independent of risk strata. As I have noted above I strongly recommend that this be revised to reflect the data I have sited. This is especially important during a time when we have rolling BCG shortages and the FDA has not moved to bring other strains through the approval process. There is ample evidence that optimized MMC significantly reduces 5 year recurrence rates compared to 20mg/20cc.	The Conclusions are consistent with the findings of the review that intravesical therapy (including MMC) reduces risk of bladder cancer recurrence, but that BCG is the only intravesical therapy shown to reduce risk of progression. It also notes that BCG is associated with a high rate of adverse events. As noted above, evidence on optimized MMC is limited to two trials with somewhat inconsistent results. As noted in the Discussion/Implications for Clinical and Policy Decisionmaking section, guidelines recommend BCG as first-line therapy, and that decisions to use intravesical therapy and choice of therapy depend on a number of factors.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4		9. Biomarker utility – I am not sure I see a discussion of the use of biomarkers to resolve atypical suspicious cytology. There is ample evidence indicating this utility, which is a common scenario encountered in the follow up of patients with NMIBC.	Use of biomarkers for evaluation of atypical cytology was outside the scope of this review, which focused on use of biomarkers for initial diagnosis and surveillance.
Peer Reviewer #4		10. Table 1 page 58 – CIS may not be an "early" event. Suggest deleting this retaining the description as a flat high grade cancer confined to the inside lining	Revised as suggested.
Peer Reviewer #4		11. Page 59 line 39 – suggest revising to use "FDA approved" only for those drugs that have an approved indication for intravesical therapy of bladder cancer.	We revised to note that electromotive administration of intravesical therapy is not widely available or in widespread use in the U.S., and not FDA approved for this purpose.
Peer Reviewer #4		12. Page 75 – throughout the review the authors refer to biomarkers studies tested in patients with "symptoms" eg FISH on this page. I find this vague as most of these trials evaluated the utility for diagnosing bladder cancer and FISH and NMP22 have specific approvals for use in this setting yet I am not sure this ever appears in this review. So it would be hard to discern this specific utility of certain biomarkers from this review.	We revised p 18 line 43-44 to be clearer that we were referring to testing for initial diagnosis of bladder cancer in patients with signs or symptoms suggestive of bladder cancer. This is to distinguish it from testing for purposes of surveillance following TURBT.
Peer Reviewer #4		13. Page 120 key points BCG. EORTC 30962 showed a benefit to standard dose vs. 1/3 dose in patients with high risk	The SOE for more prolonged versus less prolonged therapy is low

Commentator & Affiliation	Section	Comment	Response
		disease, yet the summary statement dose not reflect this. SWOG 8507 clearly shows a highly statistically significant benefit to 3 yr. maintenance vs. induction only. Why is SOE low?	because the results are based on only two trials with some methodological shortcomings. The SOE for higher versus lower doses is also low because there is some inconsistency between trials.
Peer Reviewer #4		14. Page 124 line 21-26 – see comment 7 above	The Oddens trial was included in our review; as described in the Results we believe that the optimum dosing regimen is still uncertain given inconsistent results between trials, including in higher risk patients
Peer Reviewer #4		15. Page 124 BCG strain comparisons. Need to include study by Rentsch, et al (Thallman senior author) published this year comparing induction with Connaught vs. Tice with a benefit in recurrence probability with Connaught. This is an important trial as the hypothesis is based on a specific mutation in copper	This trial has been added to the Results for the comparison of one BCG strain vs. another.
Peer Reviewer #4		16. Page 125 line 25. Trial (ref 174) also doubled the dose and doubled the concentration of MMC	We revised to make the difference in dose and concentration in the Au trial clearer.
Peer Reviewer #4		17. Table 17 page 265 – Grossman extension study reported median follow up of 53-55 months – please add this to the	We corrected this.

Commentator & Affiliation	Section	Comment	Response
		last column on right	
Peer Reviewer #5	General Comments	The identified key questions are clinically meaningful and explicitly stated, and the authors did a commendable job compiling the large number of (often heterogeneous) studies. Their summary is clear, with an excellent discussion and conclusions.	Thank you for the comment.
Peer Reviewer #5	Introduction	The Introduction provides an excellent summary of the review. It is clear and the accompanying figure (with key questions) is easy to follow.	Thank you for the comment.
Peer Reviewer #5	Methods	The inclusion and exclusion criteria are justifiable (with one comment below), and the search strategies are explicitly stated and logical, as summarized in the Methods section and detailed in the Appendix. The outcome measures are appropriate, and the statistical methods are appropriate. There are a few additional questions/comments that I have included below:	Thank you for the comment.
Peer Reviewer #5	Methods	I assume that you also excluded all meta-analyses as well (or were these used to hand-search reference lists)? I see this listed in the Literature Flow diagram in the results but not explicitly stated in the methods (and so, should be added).	We revised the Methods to be clearer that we reviewed reference lists from systematic reviews, but otherwise excluded them.
Peer Reviewer #5	Methods	For data extraction, under treatment studies (p. 10, start of line 31), why did you not include bladder cancer grade along with	We did abstract tumor grade and revised this sentence to be clearer that this was the

Commentator & Affiliation	Section	Comment	Response
		stage? This would be helpful to better understand which treatments were used in what settings (e.g. distinguishing a study that predominantly enrolled LGTa vs. HGTa- as these have different recurrence/progression profiles).	case.
Peer Reviewer #5	Methods	Minor edits: 1. p. 12, line 10. "." Between "…lesions)" and "the characteristics" should be a comma. 2. p. 12, line 51. Missing the end parenthesis.	Typos corrected.
Peer Reviewer #5	Results	The amount of detail in the results is appropriate, and study characteristics are clearly described (for the most part). Key messages are explicit & applicable. Figures, tables, and appendices are also adequate and descriptive. There is one study that was included in the references but appears to have been excluded from one of the key questions- listed below. There are also a number of comments/questions that arose during review as listed below:	This issue is addressed in KQ1a. As noted by the reviewer, accuracy increased with higher tumor stage and grade.
Peer Reviewer #5	Results	How was the grade of prior disease taken into account when providing the diagnostic accuracy of various urinary biomarkers for those undergoing surveillance. I assume that these (low and high grade) were lumped together. I think the results deserve a statement to that effect given that for some (or could argue all) biomarkers,	This issue is addressed in KQ1a. As noted by the reviewer, accuracy increased with higher tumor stage and grade.

Commentator & Affiliation	Section	Comment	Response
		sensitivity/specificity are greatly impacted by grade of disease.	
Peer Reviewer #5	Results	In Table 4, BTA qualitative vs. FISH does not have an (A) and (B) denoted- would add to improve clarity.	Typo corrected.
Peer Reviewer #5	Results	Under Head-to-Head comparisons (p.20), in the 2nd paragraph, you state that "there were no clear differences in diagnostic accuracy" for qualitative BTA vs. FISH, but the table shows a significant difference in specificity between the two (p=0.0001).	We revised the text to be clearer that there were no clear differences in sensitivity, but one study found BTA associated with lower specificity than FISH.
Peer Reviewer #5	Results	Any reason to exclude the figure for Quantitative BTA (other than the fact that only 3 studies were included)? It may be helpful just to see the 95% CI of the studies included (similar to including the figure for Immunocyt which only had 4 studies).	Because there are already so many figures, we did not add a forest plot for this analysis, which only involved 3 studies.
Peer Reviewer #5	Results	p.26, line 44- you should clarify whether the study assessed quantitative or qualitative BTA.	We revised to note that it was a study of qualitative BTA.
Peer Reviewer #5	Results	p. 38, Figure 13. It would be useful to have an explanation of the distinction of Ali-El-Dein's studies (two results from BJU, two results from JU). After scrolling to the table, I understand that these represent two different treatment groups (using different dosing regimens for Epirubicin)- but would be helpful to have a footnote denoting this in the figure (to aid the reader) given that both publications were in 1997 (and likely	We added footnotes to Figures 13 and 14 to indicate that these were different treatment regimens from the same trial.

Commentator & Affiliation	Section	Comment	Response
		represent significant overlap- which you note in the Table comments).	
Peer Reviewer #5	Results	p. 42, line 42. Please add the RR & 95% CI for progression for BCG vs. epi+IFN	The results for progression were based on survival curve estimates and not reported in a way that allowed calculation of RR and 95% CI.
Peer Reviewer #5	Results	p. 43 line 48-51- Add that this was the finding in one trial (I assume).	We revised to be clearer that the findings were based on 1 trial.
Peer Reviewer #5	Results	p. 45, line 12. 95% CI should be 0.82, 1.16 (not 0.78, 1.16) if Figure 17 is correct	We corrected this typo.
Peer Reviewer #5	Results	Figure 18- would be helpful to note difference for Ojea, 2007 (two findings- either in text or footnote). Same for figure 19.	We added footnotes to Figures 17 and 18 to be clearer that these were two different regimens from the same trial.
Peer Reviewer #5	Results	p. 57, Figure 23—The title may be incorrect or referenced incorrectly on p.45 lines 30-31--- lists as risk of recurrence. This should be risk of progression? Also, the numbers in the text (p. 45, lines 30-31 do not correspond with those numbers in Figure 23). This is confusing based on the two sections in the text: the 2nd to last paragraph under BCG vs. MMC talks about differences between BCG vs. BCG+MMC and highlights Figure 23. Yet the next section BCG+MMC vs. MMC also highlights Figure 23. It must be one or the	The Figure for BCG vs. BCG plus MMC for progression was missing and has been added. Figure 23 is labeled correctly for BCG plus MMC vs. MMC for recurrence, but will be re-numbered when the missing figure is added.

Commentator & Affiliation	Section	Comment	Response
		other since these are completely difference comparisons?	
Peer Reviewer #5	Results	p.47, lines 21-22. The first sentence/line appears to be truncated- needs to be revised.	We corrected the typo
Peer Reviewer #5	Results	p. 68-69, section "Comparisons of Different Instillation Regimens" – why was reference #80 (Lamm, 2000- SWOG trial) not included here?	This trial was included in the section on intravesical therapy vs. no intravesical therapy, but should have been included in the section comparing different instillation regimens, since all patients received induction therapy and then were randomized to maintenance versus no further therapy. We also moved the Koga and Palou trials from this section, which were designed similarly, and re-did the analysis for BCG vs. no intravesical therapy excluding these trials.
Peer Reviewer #5	Results	p. 71- line 43 references Table xx. Also, I assume "epirubicin" should be added after "intravesical"	Typo corrected. We removed the reference to a Table here as it is referred to earlier in the Results.
Peer Reviewer #5	Results	For Key Question 6, for purposes of background for the reader, it may be worth discussing that 5-ALA is not commercially available so results using HAL may be	We revised to note that 5-ALA is not commercially available. All of the forest plots are stratified by the

Commentator & Affiliation	Section	Comment	Response
		more noteworthy and clinically relevant. In saying this, it may be useful to repeat your meta-analysis, excluding those studies which use 5-ALA, and report this as a sub-analysis. This would be more clinically relevant given that 5-ALA is believed to be less effective than HAL also so using results that mix both may be misleading.	photosensitizer used (5-ALA or HAL). As noted estimates were similar regardless of the photosensitizer evaluated, so we think it is appropriate to present the analyses pooled for both photosensitizers.
Peer Reviewer #5	Results	p. 222, lines 43-44 (summary table 19)- "Two studies found sensitivity was higher for larger smaller tumors." Could you clarify this sentence?	We corrected the typo, it should just say that sensitivity was higher for larger versus smaller tumors.
Peer Reviewer #5	Results	Minor edits (I assume that the editorial process will catch most of these, but given the length of this document, I included these below in case they are helpful): 1. p.19, line 46. Missing the end parenthesis. 2. p.106, line 37 (under T1), p.106, line 47 (under T1), p.107, line 19 (under the Row T1) p.107, line 23 (under tumor grade G1)- missing a space and parenthesis so the numbers end up running together in the table. 3. Page 26, line 54. missing "a" between "find" and "difference" 4. Page 26, last line- missing comma between "prostatitis" and "benign" 5. Page 27, line 49- missing "mortality" after "all-cause" 6. Page 29-30, multiple lines- instillation is	Typos corrected.

Commentator & Affiliation	Section	Comment	Response
		misspelled as “Installation” consistently 7. Page 31, line 52. “imprecise.se” 8. Page 32, line 16. “9%% CI” instead of “95% CI” 9. Page 32, lines 23-27. Please consider dividing into two sentences for clarity. Difficult to follow otherwise. 10. Page 42, line 54- add left parenthesis before RR 11. Page 71, line 52- two other trial should be plural 12. Page 90, line 21. Remove the period preceding or hematuria 13. Page 92, line 17. Missing period at end of sentence. 14. Page 222, line 23- comparisons “were” too sparse	
Peer Reviewer #5	Discussion/ Conclusion	The implications of the major findings are clearly stated and limitations of the review (as well as included studies) are described adequately. The only potentially omitted literature is with regard to maintenance vs. induction BCG. In the results, I do not recall seeing the Lamm (ref 80) study included in the evidence with which the conclusion for SOE: low was founded. Please check the RR and 95% CI for progression under BCG vs. BCG+MMC (given the issue with Figure 23 as noted above). The future research section is clear and will	This trial was included in the section on intravesical therapy vs. no intravesical therapy, but should have been included in the section comparing different instillation regimens, since all patients received induction therapy and then were randomized to maintenance versus no further therapy. We also moved the Koga and Palou trials from this section, which were designed

Commentator & Affiliation	Section	Comment	Response
		be useful for identifying areas for which new research would be beneficial.	similarly, and re-did the analysis for BCG vs. no intravesical therapy excluding these trials. The Figure for BCG vs. BCG plus MMC for progression was missing and has been added. Figure 23 is labeled correctly for BCG plus MMC vs. MMC for recurrence, but will be re-numbered when the missing figure is added.
Peer Reviewer #5	Discussion/ Conclusion	Minor Edits: 1. Page 216, line 41. “Difference” instead of “different” 2. Page 218, line 21. Remove “a” preceding “prior” 3. Page 224, line 54- omitted “mortality” after “all-cause”	Typos corrected.
Peer Reviewer #5	Clarity and Usability	The report is well structured & nicely organized. The main points are clearly presented, with conclusions that will be useful to clinical practice and policy-makers.	Thank you for the comment.
Peer Reviewer #5	Clarity and Usability	To remain consistent with terminology (and improve clarity), exclusively using either gross or macroscopic hematuria throughout would be recommended. For example, page ES-5, line 13 “gross” and the use of “macroscopic” is noted on ES-18, line 22 and also on p.84, line 40.	We revised to use the term “macroscopic” consistently.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Clarity and Usability	Another point of clarity is to use fluorescence cystoscopy or blue light cystoscopy- and remain consistent throughout the document. For example, "blue light" is used XXX. "Fluorescent" used p.6, line 33 and in the main body of the document.	We revised to use the term "fluorescent" consistently.
Public Reviewer #1 American Urological Association	General Comments	The key outcome question related to qualitative studies of urinary biomarkers, including cytology, does not differentiate between an original diagnosis and follow-up for disease in test parameters of sensitivity and specificity. This would be very helpful for the non cystoscopist who frequently orders these tests prior to referral.	We Revised to be clearer that evaluation of symptoms was in patients without a prior diagnosis of bladder cancer. Results were presented separately for evaluation of symptoms (no prior diagnosis) and surveillance.
Public Reviewer #1 American Urological Association	General Comments	This document implies that cystoscopy is the gold standard for the identification and surveillance of urothelial malignancies within the bladder, but it does not touch on the more difficult controversy regarding the cost of urinary biomarkers, fluorescence, and the cost effectiveness of varying intervals of cystoscopy.	Assessment of cost effectiveness is outside the scope of this review, though we revised the Implications for Clinical and Policy Decision Making section to note that a factor that may impact use of these technologies is cost.
Public Reviewer #1 American Urological Association	General Comments	Likewise, on ES-37, "whether urinary biomarkers are sufficiently accurate to rule out bladder cancer… depends on the ability of clinicians to estimate pre-test probability…" Should this AHRQ report be able to spell out the pre-test probabilities? That	Reviewing the pre-test probabilities associated with different symptoms and demographic characteristics was outside the scope of this review. However, for each study on diagnostic

Commentator & Affiliation	Section	Comment	Response
		is, what are the predictive values of the presence of microscopic versus gross hematuria, irritative voiding symptoms, tobacco abuse, age, gender, etc. related to the pre-test probability of a urothelial malignancy as AHRQ has done for the progression of disease with intravesical therapies? It appears that these queries were part of the search strategies. The answer to these would be helpful in determining the answer to some of the frequent clinical questions: 1. Follow-up of the negative evaluation of microscopic hematuria related to pre-test probabilities 2. The use of urinary biomarkers in detecting urothelial disease in the de novo patient 3. The interval use of cystoscopy related to pre-test probabilities, including prior stage and grade 4. The interval use of cytology related to pre-test probabilities, including prior stage and grade 5. How often will a subsequent work up in the setting of persistent hematuria reveal a urothelial malignancy? 6. Some information for the practicing urologist on progression, stage migration, and when to refer	accuracy we describe the population included, including demographics and symptoms, and the prevalence of bladder cancer. This report does not make clinical recommendations as suggested by the reviewer; rather, it synthesizes the available evidence as determined by the key questions.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 Jim McLaughlin	Introduction	Dear Sir/Madam We have noticed that the Pacific Edge product(s) have been left out of Table 20. It is these products that we are very much most interested in as they seem to be the best performing molecular diagnostic test(s) in the market place.	We added CxBladder (the Pacific Edge product) to Table 20. Although the estimate is based on a single study, the confidence interval for the LR is similar to the other tests (which were evaluated in more studies).
Public Reviewer #2 Jim McLaughlin	Introduction	1. It is difficult to determine which molecular diagnostic test is superior from the draft report, I guess that is at the crux of what the report may provide for many. A table such as that below would be appreciated, we shouldn't have to derive our own tables. • CxBladder: Sensitivity was 0.82 specificity of 0.85 • Quantitative nuclear matrix protein 22 (NMP22): Sensitivity 0.70 specificity 0.81 • Qualitative NMP22: Sensitivity 0.58 specificity 0.88 • Qualitative BTA: Sensitivity 0.64 specificity 0.80 • FISH: Sensitivity was 0.69 specificity 0.89 • ImmunoCyt: Sensitivity was 0.77 specificity 0.77	This information is presented in Table 3.
Public Reviewer #2 Jim McLaughlin	Introduction	2. Pacific Edge are, as I understand, about to launch a second Cxbladder(triage) product and an associated study prior to the release date of the final report. Please include this	We identified no additional studies of CxBladder that met eligibility for inclusion.

Commentator & Affiliation	Section	Comment	Response
		product in the final report, I understand it may also outperform all others in respect to haematuria screening.	
Public Reviewer #2 Jim McLaughlin	Introduction	3. Some commentary on the range of clinical value propositions provided by the Pacific Edge product suite, both those in the market Cxbladder(detect) and those about to be launched Cxbladder(triage) and Cxbladder(prognosis), would be appreciated, we are very interested in these as they seem to offer both clinical and commercial benefits beyond all other molecular diagnostic tests in the market. These being; a) Replace the need for other urine-based tests in primary workup. b) Complement cystoscopy for bladder cancer detection. c) Detect urothelial tumours not visible by cystoscopy. d) Replace the need for CT / IVP in primary workup in some instances. e) Improve patient compliance with accurate, non-invasive testing. f) Complement cystoscopy for monitoring bladder cancer recurrence. g) Increase the interval between surveillance cystoscopies in certain	Only one study of CxBladder met inclusion criteria and it is presented in the report (p 20). The SOE was low based on this single (medium risk of bias) study. The purpose of this report is not to provide clinical recommendations, but to synthesize the available evidence.

Commentator & Affiliation	Section	Comment	Response
		circumstances. h) Triage patients presenting with micro-haematuria that do not need a full workup. i) Patient prioritisation in high throughput settings. j) Evaluate patients in 'at-risk' populations.	