Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: *Imaging Techniques for the Diagnosis and Staging of Hepatocellular Carcinoma*

Draft review available for public comment from December 19, 2013 to January 28, 2014.

Research Review Citation: Chou R, Cuevas C, Fu R, Devine B, Wasson N, Ginsburg A, Zakher B, Pappas M, Graham E, Sullivan S. Imaging Techniques for the Diagnosis and Staging of Hepatocellular Carcinoma. Comparative Effectiveness Review No. 143. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2012-00014-I.) AHRQ Publication No. 14(15)-EHC048-EF. Rockville, MD: Agency for Healthcare Research and Quality; October 2014. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft 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/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1990
Published Online: October 23, 2014

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Structured Abstract	[P. 5, lines 38-39] This statement needs to be qualified. The finding may be a reflection of the low pre test probability of HCC in the surveillance population. The statement needs to be carefully put in context to not mislead readers about diagnostic imaging modalities for HCC detection and their differential value.	We revised this statement to be more consistent with our findings: "Few studies evaluated diagnostic accuracy in surveillance settings (low strength of evidence), but two studies that directly compared imaging modalities found US without contrast associated with lower sensitivity than CT for detection of patients with HCC."
Peer Reviewer #8	Structured Abstract	[P. 5, lines 38-39] Another fundamental concern is the difficulty of agreeing on an acceptable reference standard. Clearly the goal of any surveillance program is the detection of a deadly disease at a curable stage. Early HCC (1-2cm) can be difficult to find by pathology correlation even if the location of the tumor is known. All of us who practice in this field know that discovery of such tumors with contrast enhanced modalities, specifically MRI is much easier than with US. In fact, many of these small nodules cannot be clearly found on correlative US especially if macronodular cirrhosis is present.	We reviewed and summarized the evidence on detection of small HCC lesions; sensitivity of all modalities was suboptimal regardless of the imaging modality evaluated.
Peer Reviewer #8	Structured Abstract	[P. 5, lines 38-39] This report should clearly state that the finding of "no difference" between contrast enhanced CT and non contrast US in detection of HCC in the surveillance population" may be as much a function of study design and inherent limitation of our ability to detect the ground truth as anything else. Obviously this statement I am making here reflects a personal bias that is based on 15 years of observation in an imaging program for a large liver clinic. I mean it to be a reality check and while the results of the literature review are what they are, the interpretation thereof should be tempered by insights from real life clinical imaging programs.	We revised this statement to be more consistent with our findings: "Few studies evaluated diagnostic accuracy in surveillance settings (low strength of evidence), but two studies that directly compared imaging modalities found US without contrast associated with lower sensitivity than CT for detection of patients with HCC." (Structured Abstract, Results)
Peer Reviewer #8	Structured Abstract	[P. 6, lines 7-8] It is not clear to me what this sentence means. What are "comparative effects on diagnostic thinking"? Please clarify	We revised the Abstract, replacing the phrase "diagnostic thinking" with "clinical decisionmaking".
Peer Reviewer #8	Structured Abstract	[P. 6, line 9] Is that an effect of the time period over which the search stretched? There have been significant advances in imaging quality in routine body MR for instance in the past 10 years. The advent of widespread use of multidetector CT in the early 2000s was another such milestone that needs to be considered. We know from the Milan study which was conducted in the first half of the 1990s that they inadvertently understaged a significant number of their patients because the CT technology at that time simply could not acquire a clean, monophasic CT of the entire liver.	As described in detail in the report, we restricted the analysis to more current technologies (e.g., multidetector or nonmultidetector CT, 1.5 or 3.0 T MRI) and performed a number of analyses based on technical parameters (e.g., number of multidetector rows for CT, imaging phases, contrast rate, section thickness).
Peer Reviewer #3	Executive Summary	P 13 not all HCC is aggressive. The doubling rate, phenotype of HCC is variable from well differentiated tumors that are incapsulated to poor differentiated tumors that are infiltrative. This is an important distinction because screening for HCC may not be beneficial for the different types of HCC.	We revised this to be clearer that the natural history of HCC is variable (ES-1, line 14).

Source: http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1990
Published Online: October 23, 2014

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Executive Summary	Is diagnostic thinking the same as clinical decision making? If not explain diagnostic thinking-p 30 line 20	We revised this to be clearer that diagnostic thinking refers to clinical decisionmaking (e.g., subsequent diagnostic testing and treatment decisions).
Peer Reviewer #6	Executive Summary	ES8 - KQ1 section - a bit confusing. Do the results refer to the 6 surveillance studies or the 174 non-surveillance studies. Hard to imagine PET being used for screening/surveillance. If non-surveillance results are reported here, then that is at odds with the question.	We revised this to be clearer that KQ 1 addresses detection of HCC in surveillance and non-surveillance settings. Evidence from surveillance and non-surveillance settings is reviewed separately.
Peer Reviewer #6	Executive Summary	ES9 (and corresponding section of main report) - First paragraph key findings, third sentence is awkwardly phrased.	Thank you. We revised this to be clearer in the Executive Summary and in the Discussion section of the report.
Peer Reviewer #6	Executive Summary	ES12 (and corresponding section in main report) - first paragraph, second and third sentences - This could be misinterpreted as suggesting that, given a dearth of evidence, AASLD recs in support of screening using U/S should be based on diagnostic accuracy studies and, in this case, this evidence supports the recommendation to screen with ultrasound. What I think you're trying to say is that, if one chooses to screen, ultrasound is probably a reasonable imaging modality to use. I think this needs to be clarified given the controversies surrounding the efficacy of screening overall. I don't think you are saying the evidence here supports screening in general (but many would seize upon the statement as suggesting it does).	We revised this to state: "Current guidelines from the AASLD recommend US without contrast for surveillance of HCC in at-risk individuals. (Bruix, 2011). Evidence from true surveillance settings to evaluate the comparative test performance of different imaging modalities was very limited. Based primarily on studies conducted in non-surveillance settings, our study suggests that US without contrast is less sensitive than MRI or CT for detecting HCC. However, findings may not be directly applicable to clinical and policy decisions related to surveillance, as the spectrum of patients evaluated in these studies could have affected estimates. In addition, decisions regarding choice of diagnostic tests to use in surveillance may depend on factors other than diagnostic testing accuracy, including costs." (Executive Summary, p.ES-23 and corresponding section of the main report, p.121).
Peer Reviewer #8	Executive Summary	[P. 13, line 35] I think it is probably critical to distinguish between contrast enhanced US (CEUS) and Ultrasound. The former has at least a chance of characterizing liver lesions while the latter essentially does not. There are regional variations on the availability of US contrast agents. In the US they are not widely available while they are commonly used in Europe and Asia. These differences need to be carefully considered and there cannot really be a "bucket" for US test performance as a whole, since these two prototypical ways in which the exam is carried out are so fundamentally different in their inherent ability to detect and characterize liver lesions.	Evidence for US without contrast and US with contrast is presented separately throughout the report. For surveillance, almost all data were on use of US without contrast, for the reasons noted by the reviewer (and noted in Table A, Executive Summary, Table 23 of the main report, and elsewhere in the report with regard to the inherent limitations of contrast-enhanced studies, in terms of limited timeframe to perform the exam, and greater usefulness to evaluate/characterize previously identified lesions).

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Executive Summary	[P. 17, lines 36-37] Does this exclusion pertain to CEUS (see comment above)? Since US contrast is not widely available nor used in the US, CEUS studies should be excluded from this review if indeed the results are to be transferred to US practice. Or, they could be included, but would need to be marked clearly as such. Just because CEUS is not currently used much in the US does not mean that should not be the case if it were clinically very useful.	In discussions with our Technical Expert Panel, we were informed that US with contrast is common practice in Canada, Europe, and Asia, efforts to obtain FDA approval in the U.S. are ongoing, and that some centers do perform (off-label) contrast-enhanced US. We therefore felt that it was clinically relevant to include contrast-enhanced-US. We revised the "Study Selection" section to be more specific that that we excluded studies of CT arterial portography and CT hepatic angiography.
Peer Reviewer #8	Executive Summary	[P. 20, lines 27-28] Per comment above, MDCT proliferated beginning in the late 90s early 2000s, so I am wondering whether all studies done before 2003 were done on comparable CT equipment (or MR equipment)	We performed additional analyses based on technical parameters, as described in the report.
Peer Reviewer #8	Executive Summary	[P. 21, lines 28-30] It would be particularly important to consider imaging performed in "surveillance" situation in tertiary referral centers critically. Often times, tertiary care settings and local hepatology centers of excellence serve as referral destinations for patients uncovered to have a "liver lesion" in local community settings. I addition those centers may also run true surveillance programs on their cirrhotic patients. At the end of the day, the liver imaging mix for studies done with the primary aim of HCC detection in those centers represents a mix of true surveillance patients and those patients who have already been diagnosed with an unspecified focal liver abnormality. The latter will have a much higher pretest probability of having the disease and therefore might significantly bias the findings derived from such centers. Just something to be very aware of and consider throughout this writeup where appropriate.	In the surveillance studies included in our report, patients were excluded if they had a previously identified liver lesion.
Peer Reviewer #8	Executive Summary	[P. 21, lines 32-33] Shouldn't these be excluded on the basis of CEUS not being performed in the US much? See comments above. I guess you will have to decide whether to systematically eliminate or exclude those studies.	In discussions with our Technical Expert Panel, we were informed that US with contrast is common practice in Canada, Europe, and Asia, efforts to obtain FDA approval in the U.S. are ongoing, and that some centers do perform (off-label) contrast-enhanced US. We therefore felt that it was clinically relevant to include contrast-enhanced-US.
Peer Reviewer #8	Executive Summary	[P. 22, lines 23-25] Across imaging modalities, sensitivity was markedly lower for HCC lesions <2 cm versus those >2 cm (differences in sensitivity ranged from 0.30 to 0.39), and further declined for lesions <10 mm in diameter	This comment is a sentence from the Discussion with no specific issue identified.
Peer Reviewer #8	Executive Summary	[P. 24, line 53] You could consider mentioning that ACRIN 6690 is under way which will shed some light on the differential performance of MR/CT http://www.acrin.org/TabID/679/Default.aspx	This is a study comparing the diagnostic accuracy of CT versus MRI in patients undergoing liver transplant; as it is similar to a number of other studies previously already included in the review, we do not think it warrants specific mention.

Source: http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1990
Published Online: October 23, 2014

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Introduction	Appropriate	Noted, thank you.
Peer Reviewer #2	Introduction	Adequate	Noted, thank you.
Peer Reviewer #3	Introduction	The introduction is well written.	Thank you.
Peer Reviewer #3	Introduction	P 25 research gaps - The authors should state that a limitation of the Zhang, et al randomized trial of ultrasound vs. no screening for HCC is the population had more hepatitis B patients then would be expected in a Western population. Hepatitis B patients develop HCC in the absence of cirrhosis and would be more likely to be resection candidates compared to a Western population where more patients with HCC have cirrhosis and are less likely to be resection candidates. Because resection is considered a cure in some patients with HCC the Zhang study may overestimate a survival benefit in a Western population.	We revised the Discussion to state: "The trial primarily enrolled patients with HBV infection, who are more likely to develop HCC in the absence of cirrhosis and therefore more likely to be candidates for surgical resection, potentially overestimating survival benefits compared to a United States population." (p.108)
Peer Reviewer #4	Introduction	Introduction makes an excellent and comprehensive case for the review.	Thank you.
Peer Reviewer #5	Introduction	The introduction provides the necessary background to set the stage for the importance of, and need for, the systematic review	Thank you.
Peer Reviewer #6	Introduction	No comments	Noted.
Peer Reviewer #7	Introduction	The introduction is BRIEF but still accurate. I personally feel that it would have benefited from some descriptions of hepatocarcinogenesis and the development of neovascularity in nodules which become malignant. This is just so important for the changes we see on imaging. I always recall how much better I performed myself once I realized the extensive variations in vascularity which are associated with different degrees of differentiation of HCC.	Page 2 of the Introduction notes that HCC lesions are typically hypervascular and explains why contrast is helpful for identifying HCC lesions.
Peer Reviewer #7	Introduction	If one considers the basis of good performance of all imaging modalities, what separates them, one from the other, is how well they pick up subtle vascularity changes throughout the period of enhancement. There is none of this within your document. I think that this is a negative.	Page 2 of the Introduction notes that HCC lesions are typically hypervascular and explains why contrast is helpful for identifying HCC lesions.
Donald Mitchell	Introduction	Adequate	Thank you.
Peer Reviewer #1	Methods	Appropriate	Thank you.
Peer Reviewer #2	Methods	Yes	Noted.
Peer Reviewer #3	Methods	The inclusion and exclusion criteria are appropriate.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Methods	The authors include nonsurveillance data in the surveillance section (p 83) and concede that this may not appropriately reflect an at risk population without a history of HCC. What were the specific settings of the at risk populations- post resection? Ablation?	We revised KQ 1 to be clearer that it addresses detection of HCC in surveillance and non-surveillance settings. Evidence from surveillance and non-surveillance settings is reviewed separately. Many of the detection studies in non-surveillance settings were performed in patients who underwent transplantation, other surgical or ablative treatments, or in patients who were already known to have HCC.
Peer Reviewer #3	Methods	The search strategies and terms are appropriate.	Thank you.
Peer Reviewer #3	Methods	The definitions of diagnostic criteria should be clarified for HCC.	Diagnostic criteria (reference standards) varied between trials from explanted liver examination to percutaneous or surgical biopsy to imaging plus clinical followup (or combinations of the above). Effects of different reference standards are addressed in KQ 1.a.i and KQ 2.a.i.
Peer Reviewer #3	Methods	The authors should state if the studies included in their analysis used the most recent guidelines of late arterial enhancement and portal venous washout on using cross sectional imaging as a diagnostic tool for HCC.	Pooled data were generally based on typical findings of arterial enhancement plus venous washout, though some studies did not describe the criteria used well. We performed stratified analysis for studies that used a formal confidence rating scale; in general there were no clear effects on estimates of diagnostic accuracy (see Tables 6-9).
Peer Reviewer #3	Methods	The statistical methods seem appropriate.	Thank you.
Peer Reviewer #3	Methods	P 49 what was used for assessment of underlying liver disease- MELD or CPT? Line 35	We reported cirrhosis rates as reported in the study; most studies did not report the criteria used.
Peer Reviewer #4	Methods	Methods are pretty much unassailable. Excellent presentation, typical of these AHRQ reviews.	Thank you.
Peer Reviewer #4	Methods	Can authors discuss how “diagnostic thinking” is operationalized in the studies considered? I am not familiar with this construct.	We revised the Abstract and Methods to be clearer that “diagnostic thinking” refers to clinical decisionmaking (e.g., use of diagnostic tests and treatments).
Peer Reviewer #5	Methods	The inclusion and exclusion criteria were fine, although i question the inclusion of indirect evidence because it has limited utility. I do realize that the body of direct or high quality evidence was limited however.	Noted, thank you for your comment.
Peer Reviewer #5	Methods	Systematic reviews were also excluded but then addressed later in the text. I would expect that the systematic reviews would have been included as part of the search and the results contrasted and compared with the current report - a minor point.	We reviewed the reference lists of systematic reviews to identify relevant studies and we contrasted our findings with those of systematic reviews in the Discussion.
Peer Reviewer #5	Methods	The search and definitions were fine.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Methods	I did wonder about the validity of results with the pooling of data given the limited quality of the evidence, no individual patient data, and the amount of (and reasons for) observed heterogeneity. Although this was addressed in the text, and the pooled evidence did not really drive any meaningful conclusions, I still wonder if the pooled analyses actually reflect the true estimate of effect.	We agree that heterogeneity was substantial; however, this is expected in reviews of diagnostic tests as noted by the Cochrane Diagnostic Test Accuracy review group and others. To address heterogeneity we performed many stratified/sensitivity analyses and also focused on direct (within-study) comparisons when available. The pooled results were generally robust and we think are informative for summarizing this very large body and complex body of literature.
Peer Reviewer #6	Methods	The methods are comprehensively described and are appropriate for this topic.	Thank you.
Peer Reviewer #7	Methods	I think that the Inclusion and Exclusion criteria are well defined and well thought out as well. Search strategies are very clearly defined.	Thank you.
Peer Reviewer #7	Methods	The Methods section is very good in my assessment. And I do think that the outcome measures are appropriate.	Thank you.
Peer Reviewer #7	Methods	However, the results are a heterogeneous amount of data which does not hang together well as little isolated facts are everywhere and they are hard to integrate into the overall message.	The findings are summarized in the Abstract, Summary of Evidence table, and Discussion, as well as in the Executive Summary.
Peer Reviewer #7	Methods	I am NOT a good person to comment on the appropriateness of statistical methods - my personal weakness.	Noted, thank you.
Public Reviewer Donald Mitchell	Methods	Adequate	Thank you.
Peer Reviewer #1	Results	Appropriate	Thank you.
Peer Reviewer #2	Results	It would be helpful to the reader to include a sentence or two on the role of alfa fetoprotein (AFP) combined with imaging for surveillance of HCC (KQ1). Data is provided in the Table A but not the text. Despite AASLD guidance recommended imaging only, in clinical practice serum AFP is often ordered along with imaging.	As noted in the Introduction (p 2, next to last paragraph), AFP is the most widely used serological marking for HCC surveillance, but recommended only as an adjunct to imaging due to limited sensitivity and specificity.
Peer Reviewer #2	Results	Inclusion of an analysis of the effectiveness of imaging in high risk versus low risk populations/individuals would be informative.	We stratified studies for KQ 1 according to whether they were in surveillance populations versus non-surveillance populations; the non-surveillance populations had higher prevalence of HCC (in many cases all patients had HCC). However, even the surveillance populations represented persons at high risk for HCC due to presence of cirrhosis, HBV, etc.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Results	It would be helpful to include a suggested algorithm for each of the key 3 questions based on the review. Understandably, this may be beyond the scope of the report but it would be useful as a quick reference for persons without the time to pour through the document.	The purpose of the evidence review is to synthesize the available evidence; developing algorithms or recommendations for imaging practice is outside the scope.
Peer Reviewer #3	Results	There are no major studies the authors overlooked except an important study on the limitation of contrast enhanced ultrasound in distinguishing peripheral cholangioma from HCC and the citation is provided below.	Thank you for the suggested study. We reviewed the cited study by Vilana et al (Hepatology 2010;51:2020-2029). It does not meet inclusion criteria because it only evaluated patients with cholangiocarcinoma (no HCC).
Peer Reviewer #3	Results	US with contrast limited specificity due to peripheral cholangiocarcinoma	The cited study by Vilana et al (Hepatology 2010;51:2020-2029) does not meet inclusion criteria because it only evaluated patients with cholangiocarcinoma (no HCC).
Peer Reviewer #3	Results	P 83- when comparing modalities in the nonsurveillance setting provide what setting the imaging modalities were being compare? Is there heterogeneity in settings among the nonsurveillance studies (post resection vs post ablation etc.)?	Evidence on accuracy of imaging to detect recurrence of HCC following treatment was reviewed separately from other studies of detection in nonsurveillance settings, which focused on accuracy of imaging prior to receipt of treatments.
Peer Reviewer #3	Results	P 90 Contrast enhanced ultrasound- suggest including discussion of limitations of CEUS in distinguishing HCC from peripheral cholangiocarcinoma (Vilana, Hepatology 2010;51:2020-9).	We reviewed the cited study by Vilana et al (Hepatology 2010;51:2020-2029). It does not meet inclusion criteria because it only evaluated patients with cholangiocarcinoma (no HCC).
Peer Reviewer #3	Results	P110 provide references in diagnostic thinking section.	The available studies, which were low quality, are summarized in the corresponding sections of the Results. As these studies do not provide reliable information for understanding comparative effects of imaging on diagnostic thinking, we don't think adding the references to the Discussion is necessary.
Peer Reviewer #4	Results	Appears to be comprehensive; I am not a content expert. Where decisions were made to conduct specific analyses, the rationale for doing so was duly provided. The tables are detailed but easily navigated.	Thank you.
Peer Reviewer #5	Results	The report is obviously very detailed, i think far too detailed especially given the majority of the evidence was not able to reliably inform strong conclusions for practice - I understand the process and need for detail in these reports however.	Noted. We attempted to summarize the voluminous evidence as best as we could.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #6	Results	There are so many results that I fear the bottom line findings may get lost. Where possible, I would recommend providing some sens/spec/LR ranges. For instance, in the exec summary and at the beginning of the respective results subsections, it would be useful to include a sentence or two with the key findings. Obviously, you won't be able to capture all permutations given the complexity of the evidence, but either an overall range or a range for results using patient as unit of analysis (which you could argue is going to be the more clinically relevant analysis) would be helpful.	We believe this information is already provided in the Key Points section for each KQ, where the pooled sensitivity, specificity, and LR's with 95% CI are provided for each modality and reference standard.
Peer Reviewer #6	Results	Page 20 main report - wasn't clear why LR presented for some of the modalities and not others (like CT).	LR's were presented unless no studies reported specificity (LR cannot be calculated in this case).
Peer Reviewer #7	Results	The identified Research Gaps are correct and appropriate especially with regards to surveillance procedures.	Thank you.
Peer Reviewer #7	Results	I have long felt that US should be the modality for screening but that US with CEUS is the only logical next step. Going from US screening to MR or to CT scan just does not work well. I find the results on Screening in your study to be astonishing. I believe that this is what you found, but it is astonishing nonetheless that there is not the clearest message that surveillance is essential.	The purpose of the report is not to make recommendations regarding the necessary of surveillance, but rather to synthesize the evidence on the comparative effectiveness of different imaging modalities.
Peer Reviewer #7	Results	Nonetheless, the key messages are explicit and applicable. The tables and appendices are certainly adequate, in fact, more so than needed.	Noted, thank you.
Peer Reviewer #7	Results	Data is also lacking - obviously because of the inclusion of data which does not provide appropriate answers. For example, in the diagnosis of HCC, BMI, Fatty Liver and Depth of lesions in the liver are potentially limiting to the success of US. For MRI, patient breathing compromises performance of a significant numbers of scans. These factors allow for proper selection but only if these factors are known. Although your study design intends to extract this type of information, seemingly, it did not.	As shown in Table 6, in studies of US that evaluated effects of lesion depth or BMI, there was no effect on sensitivity, though few studies evaluated this factor. Fatty liver was not evaluated as a potential modifier of diagnostic accuracy. Studies also did not report on the ability of patients to hold their breath; we abstracted information about scan time for MRI but were unable to evaluate this as a potential modifier of diagnostic accuracy since this information was not reported by many studies, and when reported was presented inconsistently.
Peer Reviewer #7	Results	I did consider many manuscripts I would have liked to be included that were not; however, it is difficult to have good critical analysis of how such a manuscript might meet your criteria.	Noted, thank you.
Peer Reviewer #7	Results	I find the information on screening to be astonishing. I just cannot believe that the value of screening is not CLEAR. I do myself always quote the single Chinese study to which you refer claiming reduced mortality.	The Chinese RCT that the reviewer is referring to is the only RCT that met inclusion criteria, and as described in the Results had a number of important limitations.
Donald Mitchell	Results	Adequate	Thank you.
Peer Reviewer #1	Discussion/ Conclusion	Appropriate	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Discussion/ Conclusion	Adequate	Thank you.
Peer Reviewer #3	Discussion/ Conclusion	The authors accurately describe the limited data on screening and surveillance for HCC, but the conclusions do not reflect their analysis. Given the lack of high quality randomized data to support screening/surveillance for HCC in high risk populations without a history of HCC the authors should concede that screening has become standard of care despite high quality data to support it and a randomized trial of screening versus no screening is impractical because screening has become standard of care. The authors do state that a trial comparing different imaging techniques might be feasible. Perhaps a section summarizing cost-effectiveness analyses comparing ultrasound, CT and MRI would be useful.	The evidence (or lack thereof) showing effects of screening on clinical outcomes is presented in the Results and Discussion sections. We revised the Discussion to discuss this in more depth. An evidence review funded by the VA was recently completed and addresses issues around screening in more detail. Evaluation of cost-effectiveness was outside the scope of this review.
Peer Reviewer #3	Discussion/ Conclusion	The section should discuss the diagnostic imaging criteria used among studies for diagnosing HCC on CT or MRI. How many studies use the recommended criteria in a cirrhotic liver of late arterial enhancement with portal venous washout or other imaging criteria such as a vascular blush. A subgroup analysis of the diagnostic performance of studies that used the currently accepted criteria of arterial enhancement with portal venous washout should be conducted and these results compared to studies that did not use the criteria.	We excluded data from studies that did not use standard criteria for diagnosis of HCC (arterial enhancement plus venous washout).
Peer Reviewer #3	Discussion/ Conclusion	It would useful for the authors to comment on the various staging systems for HCCAJCC/TNM, BCLC, and Milan criteria. Many of the studies evaluating imaging for staging HCC have used the TNM stage, but this is not the staging system used in clinical practice. In clinical practice the BCLC staging system is used or the Milan criteria for transplant candidates. The BCLC staging incorporates clinical factors as well. A Discussion on the complex issue of staging HCC and [remainder of comment missing]	Staging systems for HCC are discussed in the Introduction (p.1, last paragraph).
Peer Reviewer #3	Discussion/ Conclusion	P151 line 138- Did the authors mean to say ultrasound with contrast did not perform better than ultrasound with contrast?	Revised to state, “Ultrasound with contrast did not perform better than ultrasound without contrast for detection of HCC (low strength of evidence).” (p. 106, paragraph 4)

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Discussion/ Conclusion	P 154 line 54- The authors state their findings support AASLD recommendations for screening for HCC with US when they claim the one randomized study is biased and performed in China and results may not be generalizable to a U.S. population.	We have revised this as follows: "Current guidelines from the AASLD recommend US without contrast for surveillance of HCC in at-risk individuals (Bruix, 2011). Evidence from true surveillance settings to evaluate the comparative test performance of different imaging modalities was very limited. Based primarily on studies conducted in nonsurveillance settings, our study suggests that US without contrast is less sensitive than MRI or CT for detecting HCC. However, findings may not be directly applicable to clinical and policy decisions related to surveillance, as the spectrum of patients evaluated in these studies could have affected estimates of sensitivity. In addition, decisions regarding choice of diagnostic tests to use in surveillance may depend on factors other than diagnostic testing accuracy (e.g., costs) and the weight placed on any gains in sensitivity." (Executive Summary, p.ES-23 and corresponding section of the main report, p.121).
Peer Reviewer #3	Discussion/ Conclusion	They also state that data in the true surveillance setting are lacking and surveillance studies have been performed in populations likely to have a higher risk or prevalence of HCC compared to an at risk group never diagnosed with HCC. The authors need to justify this statement or concede screening has become standard of care without optimal data.	We have revised this as follows: "Current guidelines from the AASLD recommend US without contrast for surveillance of HCC in at-risk individuals (Bruix, 2011). Evidence from true surveillance settings to evaluate the comparative test performance of different imaging modalities was very limited. Based primarily on studies conducted in nonsurveillance settings, our study suggests that US without contrast is less sensitive than MRI or CT for detecting HCC. However, findings may not be directly applicable to clinical and policy decisions related to surveillance, as the spectrum of patients evaluated in these studies could have affected estimates of sensitivity. In addition, decisions regarding choice of diagnostic tests to use in surveillance may depend on factors other than diagnostic testing accuracy (e.g., costs) and the weight placed on any gains in sensitivity." (Executive Summary, p.ES-23 and corresponding section of the main report, p.121).

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Discussion/ Conclusion	P 157- conclusion- analysis does not support statement that several imaging modalities for surveillance (ongoing screening) of HCC. The results show that none of the imaging modalities have been well studied for screening and surveillance for HCC, but ultrasound has been the most studied modality and is the best option based on available data.	We have revised this as follows: "Current guidelines from the AASLD recommend US without contrast for surveillance of HCC in at-risk individuals (Bruix, 2011). Evidence from true surveillance settings to evaluate the comparative test performance of different imaging modalities was very limited. Based primarily on studies conducted in nonsurveillance settings, our study suggests that US without contrast is less sensitive than MRI or CT for detecting HCC. However, findings may not be directly applicable to clinical and policy decisions related to surveillance, as the spectrum of patients evaluated in these studies could have affected estimates of sensitivity. In addition, decisions regarding choice of diagnostic tests to use in surveillance may depend on factors other than diagnostic testing accuracy (e.g., costs) and the weight placed on any gains in sensitivity." (Executive Summary, p.ES-23 and corresponding section of the main report, p.121).
Peer Reviewer #4	Discussion/ Conclusion	This section was a bit confusing, but maybe I'm confused as to the purpose of these evidence reviews. They are not, as I understand things, practice guidelines. But the authors suggest some implications of the results for clinical practice. In most cases, these are appropriately qualified by such terms or phrases as the evidence "suggests" or the authors highlight the limited evidence on which the implication for practice is based. That said, I think some guideline developers would be very uncomfortable venturing recommendations based on this evidence, even if they invoked the "best clinical opinion" of an expert panel.	The AHRQ template for evidence reports includes a section on Implications for Clinical and Policy Decisionmaking. The report does not make recommendations for practice; rather we attempted to interpret our findings in the context of current issues relevant for clinical and policy decisions.
Peer Reviewer #4	Discussion/ Conclusion	I was missing the analogue to the strength of evidence ratings for these various interpretations, something like a "strength of [confidence in] the interpretation or recommendation for practice." It seems risky to offer any implication for practice that is prefaced with the phrase "limited evidence suggests."	The Strength of Evidence ratings are provided in the Summary of Evidence Table. Implications for practice are not rated for Strength of Evidence.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	Discussion/ Conclusion	There seems to be a disconnect between these sentences/points: (a) "Despite limited evidence in true surveillance settings, our study support [sic] current recommendations from the AASLD for US without contrast for surveillance of HCC in at-risk populations;" (p. 118) and (b) "The only randomized trial of effects of surveillance for HCC with imaging on clinical outcomes had important methodological shortcomings and was performed in China, potentially limiting applicability to screening in the United States." (p. 158). Or maybe I'm missing the point!	We have revised this as follows: "Current guidelines from the AASLD recommend US without contrast for surveillance of HCC in at-risk individuals (Bruix, 2011). Evidence from true surveillance settings to evaluate the comparative test performance of different imaging modalities was very limited. Based primarily on studies conducted in nonsurveillance settings, our study suggests that US without contrast is less sensitive than MRI or CT for detecting HCC. However, findings may not be directly applicable to clinical and policy decisions related to surveillance, as the spectrum of patients evaluated in these studies could have affected estimates of sensitivity. In addition, decisions regarding choice of diagnostic tests to use in surveillance may depend on factors other than diagnostic testing accuracy (e.g., costs) and the weight placed on any gains in sensitivity." (Executive Summary, p.ES-23 and corresponding section of the main report, p.121).
Peer Reviewer #5	Discussion/ Conclusion	I found the discussion and limitations to be thoughtfully addressed and the link between the evidence, the limitations, and the conclusions was clearly established.	Thank you.
Peer Reviewer #6	Discussion/ Conclusion	Again - would be useful to have some basic synthesis if at all possible up front in discussion. For example, you could say (I think) that for surveillance, ultrasound has moderately good sensitivity and performed similarly to CT. There are caveats to these types of statements, of course, but you could temper the confidence in the statement by giving the strength of evidence ("We found low strength evidence that ultrasound has a ___ sensitivity and performed similarly to CT in surveillance settings).	We believe the Discussion synthesizes the evidence. The strength of evidence ratings are provided in the Summary of Evidence table.
Peer Reviewer #6	Discussion/ Conclusion	page 118 -implications - Could consider including some basic clinical scenarios based on AASLD recs with LR findings. How does a negative ultrasound affect post-test probability? Could use incidence rates of HCC in say HCV cirrhotics as pre-test probability.	We added Table 24 with post-test probability for various imaging tests in populations with different pre-test probabilities of HCC.
Peer Reviewer #7	Discussion/ Conclusion	The MAJOR implication of the document is clearly stated and highlights that all modalities could be chosen for many of the aspects of HCC imaging. This is a very important conclusion.	Thank you.
Peer Reviewer #7	Discussion/ Conclusion	In my final analysis, I found the front end of this document - the introduction and objectives to be reasonable and, in fact very good. The results of the manuscript review, however, just does not give a clear message and there is just so much cluttering material that the message is obscured. I suspect that, in spite of the good intentions of the study design, the reviewed documents did not allow for clear messages to emerge.	Noted. We did our best to summarize a large body of evidence that was also very technical and had a number of important methodological shortcomings.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer Donald Mitchell	Discussion/ Conclusion	This report is based on objective evaluation of existing literature, but some limitations of the literature are not addressed. 1. Ultrasound is highly operator dependent. It is difficult, possibly impossible, to determine to what extent this limits applicability of the literature on ultrasound for HCC surveillance to general practice.	The issue of intra-operator variability is common to many medical interventions. We abstracted information about US operator from the studies when available (see Appendix I, Evidence Tables.)
Public Reviewer Donald Mitchell	Discussion/ Conclusion	2. Most literature does not define the criteria for diagnosis, so the accuracy between different reports is difficult to compare.	All studies based diagnosis of HCC on findings on contrast-enhanced imaging. We abstracted the diagnostic criteria used in the studies, including use of confidence ratings scales and features (e.g., arterial enhancement and venous washout). Stratified analyses were performed on studies that used and didn't use confidence rating scales (there was no clear effect on estimates).
Public Reviewer Donald Mitchell	Discussion/ Conclusion	3. Most literature does not define issues of confidence, e.g. probable vs definite, or probably benign vs probably malignant. However, management does incorporate this.	A number of studies utilized confidence ratings scales and this information was abstracted; however, the use of such scales and variability in the scales (e.g. 1-5, 1-4, 1-3, or other) and the cutoffs applied for a positive diagnosis varied and did not lend themselves to further analysis.
Public Reviewer Donald Mitchell	Discussion/ Conclusion	4. Two new systems for scoring CT and MRI, UNOS-OPTN and LI-RADS, include measures of confidence and more specific criteria for diagnosis. These should be discussed, including the potential impact for more standardization.	We revised the Research Gaps section to note the recent introduction of the LI-RADS and UNOS-OPTN systems and the need for additional research to determine effects on estimates of diagnostic accuracy and to identify additional opportunities for standardization.
Peer Reviewer #3	Figures	There is a large amount of detail in the Tables and Figures. The characteristics of the studies are included but further description on the imaging criteria used in the studies should be provided. The key measures are clearly stated. Figures and Tables are detailed but could be condensed.	We believe the Tables provide the critical information regarding imaging criteria; the studies used standard criteria based on enhancement patterns for HCC diagnosis. We attempted to limit the number of Figures and to shorten them as much as possible.
Peer Reviewer #3	Figures	Although I commend the authors for the very comprehensive and detailed Tables some are not useable, such as Table 7 which is 3 pages long and more than 50 rows.	The information in the Table reflects the many technical, patient-related, and methodological factors involved in assessing diagnostic performance.
Peer Reviewer #3	Figures	It would be useful to cite the references in Table 1A.	We felt addition of references would make the table difficult to read.
Peer Reviewer #3	Figures	P 32 line 34-35 table 2a should say 0.49 for lesions <2cm under ultrasound	Typo corrected in Summary of Evidence Table (p 128).
Peer Reviewer #3	Figures	P 33 table 3a- would provide gold standard under summary	We revised to be clearer that staging was based on TNM or BCLC criteria.
Peer Reviewer #3	Figures	Figure 5- p 62 why were only those 3 studies included in the Figure and not others like Zhang (randomized) study?	The table is on test performance; the Zhang study does not report test performance (sensitivity/specificity etc).

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Figures	P 93 Figure 23 is this for ultrasound with or without contrast?	It was for US with contrast, and this has been corrected. Thank you.
Public Reviewer Donald Mitchell	Figures	Adequate	Thank you.
Public Reviewer Donald Mitchell	Tables	Adequate	Thank you.
Public Reviewer Donald Mitchell	Abbreviations and Acronyms	Adequate	Thank you.
Peer Reviewer #3	References	Vilana, Hepatology 2010;51:2020-9	We reviewed the cited study by Vilana et al (Hepatology 2010;51:2020-2029). It does not meet inclusion criteria because it only evaluated patients with cholangiocarcinoma (no HCC).
Public Reviewer Donald Mitchell	References	Adequate	Thank you.
Public Reviewer Donald Mitchell	Appendix	Adequate	Thank you.
Peer Reviewer #1	Clarity and Usability	Very good, except for the point expressed above	Noted, thank you.
Peer Reviewer #2	Clarity and Usability	Yes	Thank you.
Peer Reviewer #3	Clarity and Usability	The study is well written.	Thank you.
Peer Reviewer #3	Clarity and Usability	It could be better organized. The first 36 pages are an executive summary and this should be shorter.	We wrote the Executive Summary in accordance with current AHRQ guidance on elements to be included.
Peer Reviewer #3	Clarity and Usability	This is followed by the manuscript which includes the same Figures as the executive summary- Figure A p15, Figure B p 16, Figure C page 16 are the same as Figure 1, 2 page 41 and Figure 3 p 42. There are other examples of redundancy.	The Executive Summary is meant to be a standalone document so some redundancy in tables and figures is intentional.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	Clarity and Usability	Besides the points I raised above, I think the report is well organized and could be used by policy makers/guideline developers as the basis for policy and/or practice recommendations. These groups would need to make their own judgments about the value of the often limited evidence (and almost complete lack of data on adverse effects) for concrete recommendations.	Noted, thank you.
Peer Reviewer #4	Clarity and Usability	To beat a dead horse, I think an appropriate conclusion of an evidence review by AHRQ is (for instance) that the evidence on "x" is sparse and not of high quality as conventionally construed, assessed. Let the policy makers translate the evidence into practice. The presentation of the evidence, as structured, would allow this. I will stop now. Not sure exactly what the evidence review group was asked to do.	Noted. The evidence review synthesizes the evidence and does not make recommendations for practice.
Peer Reviewer #5	Clarity and Usability	The report is well structured and organized - too long in my estimation, but well done. This report could be used as the basis for practice decisions. Even though I have some doubts about the quality of the evidence, pooled or not, to inform conclusions, the conclusions that were reached accounted for the weakness of the data and are therefore appropriate in my opinion.	Thank you for your comments.
Peer Reviewer #6	Clarity and Usability	See points above. Overall, well done report. A few summary statements in strategic points through the report could be very helpful in improving the clarity/usability.	Noted, thank you.
Peer Reviewer #7	Clarity and Usability	The report is certainly well structured and reasonably well organized. The report is clear in terms of the main objective. This message is successfully relayed.	Thank you.
Peer Reviewer #7	Clarity and Usability	However, with regards to secondary objectives, the result is less well defined. Further, there are conclusions in the manuscript that surprise me, for example, lack of good documentation on the value of screening for HCC.	Noted. The evidence on clinical outcomes associated with imaging for screening is presented in the report (see KQ1.c).
Peer Reviewer #7	Clarity and Usability	In real practice, I think that the value of each modality is absolutely clear. However, there are just so many important decisions related to imaging for HCC that are just not even touched upon in this document.	Noted. The evidence review synthesizes the evidence and does not make recommendations for practice.
Peer Reviewer #8	Clarity and Usability	See general comment above. Complete separation throughout the text of the surveillance and the diagnostic/characterization of focal liver lesion situations should be attempted. The available underlying evidence, results and clinical implications of these two distinct clinical scenarios are strikingly different and must not be mixed up. The report accomplishes this task to some degree in the breakout of the key questions, but in the remainder of the text the separation is less clear.	Evidence for detection (KQ 1) in surveillance and nonsurveillance settings and for evaluation of previously identified lesions (KQ 2) is reported and summarized separately.
Peer Reviewer #1	General	The analysis appropriately makes a distinction between surveillance and diagnosis. In the surveillance setting it is not necessary to correctly identify a lesion as HCC, only to demonstrate that the lesion is present. This is why non-contrast ultrasound works in surveillance, but not in diagnosis. Although this is implied in the discussion it is not explicitly stated	Limitations of contrast-enhanced US for surveillance are discussed in Table 1 and the Introduction.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	General	Yes	Noted.
Peer Reviewer #3	General	Graham E, et al. performed an extensive, comprehensive analysis of the existing literature on ultrasound, CT, MRI and PET scan for surveillance, diagnosis and staging of HCC. The analysis is well performed and summarizes qualitatively and quantitatively studies on the topic.	Thank you.
Peer Reviewer #3	General	The final conclusions are not supported by their analysis and should be modified. There is redundancy throughout the manuscript and if the authors remove the redundancy then the manuscript could be more concise and shortened.	We attempted to summarize this very voluminous and technical body of literature as much as possible and edited the text to reduce redundancy where we could.
Peer Reviewer #3	General	I would suggest including discussion on the performance of bone scan for detecting metastatic disease.	Bone scan was not an included imaging modality for this report.
Peer Reviewer #4	General	The report is clear. Key questions and corresponding analytic frameworks are presented with pristine clarity. Tables are exhaustive and easy to read. My comments focus on the authors' interpretation of the results.	Thank you.
Peer Reviewer #5	General	The report addresses an important topic with key questions clearly stated	Thank you.
Peer Reviewer #6	General	This is a comprehensive, methodologically sound systematic review of a topic that is very clinically important and is relevant to primary care physicians, radiologists, and hepatologists. This represents an enormous undertaking - a huge number of studies of different tests and conducted in different settings. The population and questions are well-described and defined.	Thank you for your comments.
Peer Reviewer #7	General	YES, the target population and audience are explicitly defined. And the key questions are explicitly stated. In fact, The KEY QUESTIONS are great and would be standard for anyone doing imaging considering the value of different techniques.	Thank you.
Peer Reviewer #7	General	However, the questions about the different reference standards, intermediate outcomes, patient outcomes and adverse effects are difficult in that I cannot even imagine the response or its value.	The Key Questions and outcomes were developed based on the original topic nomination and with input from a Technical Expert Panel.
Peer Reviewer #7	General	I extract the following from the document: "The information in this report is intended to help health care decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well informed decisions and thereby improve the quality of health care services." This is a very credible statement and there is no doubt that in terms of evidence based research the document does convincingly show that US, CT and MR are all effective imaging techniques for HCC for test performance at least.	Noted, thank you.
Peer Reviewer #7	General	The success in terms of the other objectives of diagnostic thinking, clinical outcomes and harms is much less clear and I did not come away with any message in this regard.	As noted in the report, evidence for these areas was very limited.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #7	General	The target audience is very broad and includes people of different backgrounds and motivations. To make a document with equal appeal across this broad spectrum of people is difficult I believe. Furthermore, although it is incredibly valuable to everyone in your target audience to know that all of the three major cross sectional imaging modalities, US, CT and MR have similar performance for diagnosis of HCC, as a clinician, showing the sensitivities and specificities for HCC diagnosis in so many studies is not so appealing to me. Alternately, I have huge interest in the specifics of diagnosis on all of the modalities which is obviously not the focus of a document such as this.	The report is intended to be useful to a broad variety of stakeholders. We focused on the outcomes as defined in the PICOTS and Key Questions.
Peer Reviewer #7	General	Regarding the decisions that might be made on the basis of evidence based research, I do believe that this has huge value for policy makers and health care decision makers although I do not believe that such information will in fact have value for patients nor in fact for most clinicians beyond showing that all of the modalities have the potential to do a good job. There is however, very good information on the influence for diagnosis of lesion size and also degree of tumor differentiation. This type of information is very essential and it is one of the areas of greatest interest to me in this effort.	Noted, thank you for your comments.
Peer Reviewer #7	General	Although the authors do seemingly meet their major objective and show effectively the different performances of so many modalities in so many publications, the inclusion of the PICOTS is just rather distracting as there is just no uniformity of response or appropriateness.	The PICOTS are important for defining the scope of the review and are standard for AHRQ-funded comparative effectiveness reviews.
Peer Reviewer #7	General	In my own practice, I have a very strong belief and awareness of screening ultrasound (US), diagnostic US, and US for surveillance of those with HCC. I know all of the strengths and weaknesses. From this document, I do not think that any person could make good and appropriate decisions about US at all. My thoughts in this regard are aggravated by the information about the lack of an American bubble approval and even inclusion of the disastrous and unreasonable black box warning by the FDA in 2008, which does not, in my wildest thoughts, have any relevance to this manuscript.	Noted. The evidence regarding US with contrast as well as without contrast is presented in the report. Although contrast agents for US are currently not approved by the FDA for use in liver imaging, efforts to obtain approval are ongoing and these agents are used in many other parts of the world.
Peer Reviewer #8	General	Key questions appropriate and explicitly stated.	Thank you.
Peer Reviewer #8	General	Target population well defined but heterogenous (surveillance versus focal liver abnormality). Authors attempt to separate the two, but in the flow of the text that separation is not always clearly accomplished. Perhaps an even more stringent effort of separation needs to be made. Underpinning those 2 populations, data and results are very different.	Results for detection (in surveillance and non-surveillance settings) and evaluation of previously identified lesions are presented separately.
Peer Reviewer #8	General	Authors should mention on going trials, currently the review is only retrospective. The research gaps section may offer an appropriate venue for that.	We searched for ongoing RCT's of surveillance through searches on clinical trial registries but found none.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	General	While the gist of the report should reflect the findings of the literature review, I would strongly suggest that the interpretation of the results be a bit more guided by real world clinical observations/realities. For instance, to state that no difference was found between the ability of non contrast US and contrast enhanced CT to detect HCC in a surveillance population may be a true statement based on the review, but anyone practicing clinical imaging in this space knows that it is not a realistic representation of what we see in practice. It is a result, which in my opinion is generated by a combination of unsuitable underlying study designs and more importantly a general lack of a credible ground truth in this early HCC detection space. Authors should carefully consider how to put this into context for the purpose of the report. If this report is to inform clinical guidelines, in this particular context it will serve to extend a situation where the guideline may recommend one thing (US screening) and almost everyone in clinical practice is doing something else. This truly represents a conundrum	The purpose of the guideline is to synthesize the available evidence. Results for surveillance are presented in KQ 1, including limitations in terms of the number of studies and methodological limitations in the studies.
Peer Reviewer #8	General	As with all meta analyses, the quality of the ultimate observations and responses to the key questions are held hostage to the quality (or lack thereof) of the underlying universe of studies. What is available to date in the way of rigorous studies in this HCC imaging/detection space leaves a lot to be desired. This should perhaps come out more clearly in the report as it represents a severe limitation of this entire effort. I can see why the authors may be inclined to not overemphasize this fact but it is a reality and must be said.	There is a section focusing on Limitations of the Evidence Base in the Discussion.
Public Reviewer Gail Rodriguez (MITA)	General	The Medical Imaging & Technology Alliance (MITA) is pleased to submit comments on the Agency for Healthcare Research and Quality (AHRQ) draft comparative effectiveness review entitled *Imaging Techniques for the Surveillance, Diagnosis, and Staging of Hepatocellular Carcinoma* ("Draft Report").1 MITA has extensive knowledge of the substantial benefits afforded by medical imaging and radiation therapy to the health of Americans due to our role as the leading trade association representing medical imaging, radiation therapy, and radiopharmaceutical manufacturers. We support quality efforts that foster appropriate use of these technologies for the early detection, diagnosis, staging, therapy monitoring, and surveillance of many diseases.	Thank you for your comments.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer Gail Rodriguez (MITA)	General	Medical imaging encompasses X-ray imaging, computed tomography (CT) scans, diagnostic ultrasound, nuclear imaging (including positron emission tomography (PET)), magnetic resonance imaging (MRI), and related imaging acquisitions. Medical imaging is used to diagnose patients with disease, often reducing the need for costly medical services and invasive surgical procedures. In addition, medical imaging equipment often is used to select, guide, and facilitate effective treatment, for example, by using image guidance for surgical or radiotherapeutic interventions.3 MITA's members also develop and manufacture innovative radiotherapy equipment used in cancer treatment.	Noted, thank you.
Public Reviewer Gail Rodriguez (MITA)	General	Our comments address three areas in the Draft Report: (1) imaging modalities have varied functions and uses in a clinical setting; (2) outcomes related to the use of imaging must be defined to reflect the unique contribution of imaging to clinical decisions; and (3) innovative, dose-lowering imaging technologies support quality care.	Thank you for your comments.
Public Reviewer Gail Rodriguez (MITA)	General	1) Imaging modalities have varied functions and uses in a clinical setting. As such, comparative analyses of modalities are of limited value, especially when removed from the particular clinical setting and circumstances of the individual patient.	We disagree with the contention that understanding comparative accuracy and effects of imaging is of limited value; rather such information is critical for informing diagnostic practices.
Public Reviewer Gail Rodriguez (MITA)	General	Medical imaging includes multiple modalities and each modality provides unique and many times complementary value in better understanding the clinical situation. In fact, outside the context of a particular episode of clinical care, comparisons of modalities do not appropriately value the contribution of each modality to healthcare. Rather imaging modalities should be considered in the context of the information they add to the clinical situation and how they add value in establishing appropriate care for the individual patient. AHRQ states that there are "differences in test performance between different imaging modalities and techniques" and calls for more research in this area.4 However, these stated differences in technology are of limited meaning outside the context of patient care. Currently, no single imaging technology provides all necessary information to care for every patient in every clinical setting.	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	Access to appropriate imaging is necessary to inform clinical decisions related to the proper diagnosis and treatment of disease. In order to better direct the optimal use of imaging, physician societies and other provider groups have developed appropriate use criteria and practice guidelines specific to individual clinical indications. These clinical decision-support tools are based on research and evidence, and aid physicians to determine the appropriate scans to be used for specific clinical indications.	Noted.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer Gail Rodriguez (MITA)	General	The National Comprehensive Cancer Network (NCCN) has clinical practice guidelines on hepatobiliary cancers.5 The guidelines outline considerations and approaches to care. For each stage of care, appropriate testing and treatment are outlined. In addition, imaging modalities are discussed. In fact, the guidelines state, "The number and type of imaging is dependent on the size of the liver mass or nodule." For example, liver lesions of less than 1 cm should be evaluated by 3-phase contrast-enhanced CT, MR, or ultrasound, but liver lesions that are greater than 1 cm in size should be evaluated with 3-phase contrast-enhanced CT or MR. Theseguidelines appropriately acknowledge that clinical value of each imaging modality is determined by how it informs specific clinical care, not how it ranks in comparison to other modalities.	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	MITA advocates the development and use of physician-developed appropriateness criteria to guide treatment decisions and training of hospital and imaging facility personnel who perform medical imaging exams. In order to provide optimal care and prevent medical errors, physicians and technologists must account for the patient's individual needs. By providing proper training and adhering to these standards and initiatives, physicians can ensure that patients receive the life-saving benefits of medical imaging technology.	Noted. This issue is outside the scope of this report.
Public Reviewer Gail Rodriguez (MITA)	General	2) Outcomes related to the use of imaging must be defined to reflect the unique contribution of imaging to clinical decisions.	Improvements in clinical outcomes are the ultimate goal of every medical procedure, including imaging tests, and as such are the most important outcome. In fact, studies that address clinical outcomes are placed at the top of the diagnostic evidence hierarchy by GRADE and others. Although such data are often lacking for imaging tests, there are a number of cases where this is not the case (mammography for breast cancer screening, CT for lung cancer screening, imaging in patients with LBP, etc.).
Public Reviewer Gail Rodriguez (MITA)	General	The Draft Report points to lack of studies on "clinical outcomes". This is cited as a gap in evidence. However, we offer that this is not a gap, but rather includes endpoints which are inappropriate to evaluate diagnostic imaging in the context of patient care.	Improvements in clinical outcomes are the ultimate goal of every medical procedure, including imaging tests, and as such are the most important outcome. In fact, studies that address clinical outcomes are placed at the top of the diagnostic evidence hierarchy by GRADE and others. Although such data are often lacking for imaging tests, there are a number of cases where this is not the case (mammography for breast cancer screening, CT for lung cancer screening, imaging in patients with low back pain, etc.).

Commentator & Affiliation	Section	Comment	Response
Public Reviewer Gail Rodriguez (MITA)	General	One consideration is that it is difficult to isolate the contribution of diagnostic imaging from the larger care paradigm, and in fact, due to the incremental value of diagnostic imaging within the delivery of healthcare, diagnostic imaging's value outside the care paradigm would be of limited meaning. Models that attempt to extract diagnostic imaging from the care that it informs neglect to reflect the reality of healthcare delivery. In fact, in clinical practice, a patient may have multiple diagnostic tests, with additional value from each test used to inform the complex clinical decision process in unique and inimitable ways. In addition, some diagnostics tests are synergistic. For example, a CT scan may be ordered in follow up to an ultrasound scan that shows a mass or nodule. Additionally, as the science of cancer staging progresses, diagnostic imaging may inform decision-making in concert with other tests including biomarker identification, genomic studies, and other assays.	Key Questions 2 and 3 focus on followup testing (evaluation of a previously identified nodule or staging) and we also summarized evidence on using imaging modalities in combination.
Public Reviewer Gail Rodriguez (MITA)	General	AHRQ also acknowledges this synergistic interaction between diagnostic technologies: "Understanding the diagnostic accuracy of imaging methods and how they affect clinical decision making and, ultimately, patient outcomes is a challenge. Imaging techniques may be used alone, in various combinations or algorithms, and/or with liver-specific biomarkers, resulting in many potential comparisons."	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	A more appropriate endpoint for diagnostic imaging would be similar to that which AHRQ considers as "intermediate outcomes" including effects on diagnostic thinking and clinical decision making. That is, changes in therapeutic management or stage reclassification are appropriate terminal points when considering the impact of diagnostic imaging on healthcare. A recent article on the topic suggests, "The outcomes, or endpoints, appropriate to assessing whether diagnostic interventions are reasonable and necessary are best characterized as "change in clinical management." This is distinct from the outcomes, or endpoints, classically applied in assessing whether therapeutic interventions are reasonable and necessary."	Effects on diagnostic thinking and clinical decision making were reviewed.
Public Reviewer Gail Rodriguez (MITA)	General	3) Innovative, dose-lowering imaging technologies support quality care.	Thank you for your comment. The report notes that no study was designed to assess long-term harms associated with radiation exposure; the Discussion (p 113) cites recent data from the Radiological Society of North American and the American College of Radiology regarding radiation doses with CT and PET/CT.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer Gail Rodriguez (MITA)	General	The Draft Report also points to radiation dose as a potential harm of CT and PET.12 In recent years, innovative, dose-lowering technologies have limited dose while maintaining imaging quality. Due to lower dose and high clinical efficacy, the CT and PET/CT benefit-to-risk profiles have improved.	The report notes that no study was designed to assess long-term harms associated with radiation exposure; the Discussion (p 113) cites recent data from the Radiological Society of North American and the American College of Radiology regarding radiation doses with CT and PET/CT. Though the efforts to reduce radiation exposure are laudable they do not represent "evidence" about harms.
Public Reviewer Gail Rodriguez (MITA)	General	Dose efficiency and dose reduction have been important design considerations for CT for many years. The focus on these design considerations has grown and intensified in more recent years, and has yielded a variety of new and innovative hardware and software features that directly help physicians both reduce and monitor dose for CT exams. The CT industry has developed new features that enable both the dose to be displayed prior to scanning, and to alert operators to potentially higher than expected doses, as well as enabling electronic recording of the CT dose in the patient record. These features are important for both the patient as well as facilities, since they provide facilities with the ability to compare the dose of their CT protocols and establish optimized reference values.	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	The dose monitoring/reduction features described below play a significant role in helping to reduce the dose for CT exams, while maintaining diagnostic quality and the capability to report and record dose. For example:	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	Automatic Exposure Control helps optimize dose for each patient for the given diagnostic task. This feature adjusts the exposure to use only what is needed to maintain a constant image quality. This feature is now standard on CT systems. Wider coverage detectors minimize the amount of x-ray that falls outside of the active detector region, thereby reducing dose to the patient without impacting image quality. Systems are now available in a range of wide coverage designs. "Shutter" modes block unused x-ray at the beginning and end of helical scans and therefore do not degrade image quality. This feature is now standard on many CT systems and is "built in" to each helical acquisition. Advanced electronics in data acquisition systems result in better imaging performance and less noise, thereby enabling equal performance at a lower dose.	Noted.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer Gail Rodriguez (MITA)	General	First generation CT iterative reconstruction results in a significant dose reduction potential, while maintaining diagnostic image quality, and is well suited to CTC studies. Iterative reconstruction is available on new systems and also as an upgrade to many installed base systems. More advanced second generation CT iterative reconstruction provides even further dose reduction potential, where some expert users are able to achieve some exams approaching 1 mSv levels for combined supine and prone CTC scans, while still maintaining diagnostic image quality. This feature is becoming widely available on new systems. The DICOM Dose Structured Report allows the exam dose to be electronically captured with the patient record. This feature is now standard on all new CT systems and has also been implemented on newer installed base systems.	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	MITA leads industry efforts to coordinate and establish standards to mitigate radiation dose. Adoption of these standards benefits patient dose. MITA's approach builds upon existing manufacturer safety measures – including equipment safety standards, protocol development, quality and safety checks, provider education programs and physician-developed medical guidelines – to minimize radiation dose as much as possible, and to provide even greater degrees of coordination, transparency and reporting in the delivery of medical radiation. Recent examples of MITA standards which have addressed dose include:	Noted, thank you.
Public Reviewer Gail Rodriguez (MITA)	General	NEMA XR 25-2010, *Computed Tomography Dose Check*. This standard introduced two novel features to assist the imaging team in providing better patient care: dose notifications and dose alerts. Dose notifications are designed to provide a clear indication to health care providers when the parameters for a CT scan will result in a dose higher than the facility's pre-determined dose threshold for routine use. Dose alerts are designed to prevent dose levels for a complete exam from exceeding pre-determined thresholds that are deemed excessive by the facility. This feature can be configured to prevent equipment operation. These protections help the operator and ultimately the physician to better understand dose implications of protocol choices, and should significantly reduce exposure due to inappropriate scan parameter settings.	Noted.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer Gail Rodriguez (MITA)	General	NEMA standard XR 26 - 2012, *Access Controls for Computed Tomography: Identification, Interlocks, and Logs.* This standard requires software features that ensure only an authorized operator can alter the controls of CT equipment. This industry-wide standard requires the institutionalization of administrative privileges, access levels, and the recording of clinical protocols to ensure safe and appropriate use. NEMA standard XR 27 - 2012, *X-ray Equipment for Interventional Procedures User Quality Control Mode*. This standard helps imaging facilities conduct quality testing and monitoring of X-ray equipment used for interventional procedures.	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	NEMA standard XR 29 - 2013, *Standard Attributes on Computed Tomography (CT) Equipment Related to Dose Optimization and Management*. This standard, known also as MITA "*Smart Dose"*, is the fourth dose-related standard to be released by MITA since 2010. This standard includes four components: 1. DICOM Dose Structured Reporting – This enables the recording of post-exam dose information in a standardized electronic format. This information can be included in the patient record, promoting the establishment of diagnostic reference levels, as well as facility dose management and quality assurance. 2. Pediatric and adult reference protocols – These are a set of pre-loaded protocols on a CT system that serve as a baseline for a variety of clinical tests.	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	3. CT Dose Check – CT Dose Check incorporates two features—dose notifications and dose alerts that can inform operators and physicians when dose exceeds established thresholds. 4. Automatic Exposure Control (AEC) – AEC automatically adjusts the amount of radiation used based on the size, shape and composition of the patient, in order to achieve a specified level of image quality.	Noted.
Public Reviewer Gail Rodriguez (MITA)	General	MITA appreciates this opportunity to comment on the Draft Report. We would be pleased to answer any questions you might have about these comments. Please contact me at (703) 841-3235 if MITA can be of any assistance.	Noted, thank you for your assistance.