*Research Review Disposition of Comments Report*

**Research Review Title:** *Decision Aids for Cancer Screening and Treatment*

Draft review available for public comment from July 10, 2014 to August 7, 2014.

**Research Review Citation:** Trikalinos TA, Wieland LS, Adam GP, Zgodic A, Ntzani EE. Decision Aids for Cancer Screening and Treatment. Comparative Effectiveness Review No. 145. (Prepared by the Brown Evidence-based Practice Center under Contract No. 290-2012-00012- I.) AHRQ Publication No. 15-EHC002-EF. Rockville, MD: Agency for Healthcare Research and Quality; December 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://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2029*
*Published Online: December 23, 2014*

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #1** | General Comments | Excellent description of the populations targeted and the settings for delivery. | No action required. |
| **TEP Reviewer #2** | General Comments | Generally well done and complete. I have no substantiative corrections to make. Only recommended changes are to improve understandability of portions of the text which were difficult to understand as written. | No action required. |
| **TEP Reviewer #3** | General Comments | Yes, this report is clinically meaningful and adds to Dawn Stacey's work as the target population is more homogeneous and additional studies have been found compared to her Cochrane review. It also adds to the knowledge base through the detailed analysis of effect modifiers. | No action required. |
| **TEP Reviewer #3** | General Comments | One could argue that clinical and especially shared decision making is a complex phenomenon to execute, to influence and to measure. This may therefore justify the systematic review of other than the classical epidemiological rigorous experimental designs such as RCTs. Perhaps something to reflect upon in the discussion ..... | The reviewer's point is now incorporated into the Discussion. "Decision aids are complex interventions, and their successful integration and continued use in routine care depends on many factors, including patient and provider acceptance, system infrastructure, fit with other processes, and other factors only peripherally related to the patient-provider dyad. Thus, implementation of decision aids interventions in routine practice requires consideration of many additional factors. Although we looked for studies of the effectiveness of interventions to providers for promoting shared decisionmaking through decision aids, we found limited evidence. A more general treatment of shared decisionmaking promotion interventions did not draw strong conclusions" |
| **TEP Reviewer #4** | General Comments | The authors have compiled a comprehensive systematic review to examine two key questions with regards to the effectiveness and use patient decision aids for cancer screening, prevention or treatment decisions. The key questions are well defined and target important unanswered questions in the field. The study is detailed and does present some new information that will be helpful to researchers and designers of decision aids. | No action required. |
| **Peer Reviewer #1** | General Comments | This systematic review is a commendable effort to appraise evidence on the effectiveness for decision aids related to cancer. Moreover, it attempts to answer several questions with differing methods. | No action required. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #1** | General Comments | The target population and key questions are clearly defined. However, the population is heterogeneous; on the one hand people who are screened or preventively treated for cancer, on the other patients with early cancer. | We agree with reviewer's point and this is the reason why we included a separate analysis for these populations. |
| **Peer Reviewer #1** | General Comments | Methods comprise a systematic appraisal of available studies, triangulation by a panel of stakeholders, and analysis of effect modifiers. Hence, the title of the paper seems not comprehensive. | This terminology is typical for CERs in the program and thus no change the title of the report is required. |
| **Peer Reviewer #1** | General Comments | Also, I am not yet convinced about the additional value of this review given the evidence already available from previous systematic reviews, which the authors already admit: "Other published research where such evidence for cancers was systematically reviewed reached similar conclusions." Giving more attention in the text on what is now tucked away in Appendix C may help. | This is now more explicitly addressed at the end of the introduction: "The current systematic review is designed to address issues relating to content and format of decision aids in terms of their intended audiences, as well as factors related to provider utilization." |
| **Peer Reviewer #2** | General Comments | The report is clearly reported and addresses clinically meaningful questions explicitly and thoroughly. | No action required. |
| **Peer Reviewer #3** | General Comments | This is a very carefully conducted and robust systematic review. | No action required. |
| **Peer Reviewer #4** | General Comments | This review is rigorously executed and clearly summarized. The findings are meaningful and useful. | No action required. |
| **Peer Reviewer #4** | General Comments | The key questions are appropriate and explicitly stated. | No action required. |
| **Peer Reviewer #4** | General Comments | It would be helpful to provide an explicit rationale for examining cancer risk group and literacy level as effect modifiers. | |
| **Peer Reviewer #5** | General Comments | The report addresses the impact of patient decision aids for cancer screening, prevention, and early treatment. The target conditions are reasonable. | No action required. |
| **Peer Reviewer #5** | General Comments | The key questions are interesting but a bit limited, and my sense is the findings are largely predictable. This is not to say the findings aren't important, but they largely align the systematic reviews of decision aids more broadly. | We acknowledge the reviewer's point and would like to clarify that the assessed key questions were the product of extensive discussion during the topic refinement and the review protocol process. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #5** | General Comments | I want to challenge the authors a bit about key question 1. The review is somewhat inconsistent about considering the impact of aids on behaviors/choice. They indicate choice should not be a primary outcome, yet included actual or intended choices in the review. While everyone agrees decision aids should lead to informed, values-based decisions, it is still important in my mind to consider choice given the concerns about practice variation, etc. Reworking the results to look at how the aids impacted choice for the 3 main cancers should at least be considered. For example, for prostate cancer I am concerned the review missed 2 previous reviews which each found decision aids decreased uptake of PSA testing. | We thank the reviewer for his insightful comment. During the extensive discussions with the TEP members and given the large variety of the assessed study groups, it was decided not to focus exclusively on the actual choices made as these would be very specific to the cancer type. Nevertheless, any outcomes related to actual choice were captured. Of these data, relatively few referred to how often the choice was changed before and after the DA or with/without the DA thus rendering a quantitative synthesis not informative. |
| **Peer Reviewer #6** | General Comments | This review evaluated literature looking at effectiveness of decision aids. It is generally well-conducted and here are some comments for the authors to consider. | No action required. |
| **Public Reviewer #1 (Oncology Nursing Society)** | General Comments | Key Question 1 focuses on evaluating the benefits and potential adverse effects (such as increased anxiety, etc.) on patients through the use of decision aids. It is valuable to affirm through the results of this report that the strongest evidence indicates that patients do increase their knowledge of treatment options without increasing decisional conflict or anxiety directly from interaction with the tools. These results are helpful to ONS' members In those settings Where decision Aids are routinely employed. | No action required. |
| **Public Reviewer #1 (Oncology Nursing Society)** | General Comments | In addition, the report notes that the evidence Is inconsistent when it comes to assessing the impact of decision aids in improving patient-provider communication, shared decision making, and with patient satisfaction with the decision-making process. These areas are of key interest to ONS' members, and we will communicate these findings to our research staff to inform future research priority area work. | No action required. |
| **TEP Reviewer #1** | Introduction | The authors described the impetus for this study well. | No action required. |
| **TEP Reviewer #2** | Introduction | page 10, line 15, change "...decisions about cancer screening, or treatment..." to "...decisions about cancer screening and prevention or treatment..." | We made this change using the wording suggested by the reviewer. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #3** | Introduction | The key questions are appropriate and explicitly stated although the title of the review seems to cover key question 1 only. Perhaps you chose deliberately to do so given the findings of the review. | Indeed, KQ2 is providing some additional information on promoting use of shared decision making with DAs, and is complementary to KQ1. However, it is too narrowly specified to address interventions for promotion of shared decision making in general. For this reason, it is not emphasized in the title. |
| **TEP Reviewer #3** | Introduction | I miss an explanation for your focus on screening and early stage cancer only. | In response to reviewer 2's comments, we added "The panel agreed that this population included not only patients with early cancer, but also patients who are either at high risk of cancer or are at average risk and are deciding whether to be screened. These populations could be examined in aggregate because the types of decisions being made were similarly equivocal in terms of both benefits and harms." |
| **TEP Reviewer #4** | Introduction | P 7 Abstract: the first sentence in the review methods is a bit confusing. Perhaps rewording to something like, "we included randomize controlled trials where DA interventions were compared to one or more control or intervention arms. | Rephrased using the reviewer's wording. |
| **TEP Reviewer #4** | Introduction | P 7 Abstract review methods second sentence – remove "already developed" perhaps better to focus on consumers actually making decisions? | Now rephrased to: "We included trials of previously developed DAs delivered at the point of the actual decision." |
| **TEP Reviewer #4** | Introduction | P 7 Abstract review methods: anything to add for key question 2 here? | We changed the last sentence to "For Key Question 2, we included studies of any intervention to promote decision aid use, regardless of study design and outcomes assessed." |
| **Peer Reviewer #1** | Introduction | The authors state that: "information is important for developing practical guidance about designing and using decision aids, particularly for decisions related to early cancers". How has their review provided this information? | We now further clarify in the Discussion: "This suggests that simpler (and less costly to develop and maintain, and easier to use) decision aids may be as effective as more complex ones." |
| **Peer Reviewer #2** | Introduction | Clearly written - does a good job of setting the context for the work that was undertaken. | No action required. |
| **Peer Reviewer #3** | Introduction | The report is clinically meaningful. | No action required. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #3** | Introduction | The target population and audience are not explicitly defined. | We now clearly identify this information in the Introduction:<br>Added to page 9: "…particularly for decisions related to screening, prevention, or treatment of early cancers, the target population for this review…" and "We triangulated the importance of these issues by engaging a diverse panel of stakeholders, including developers and users of decision aids, representatives of professional societies, patient advocates, and non-syndicated patients, representing the review's intended audiences." |
| **Peer Reviewer #3** | Introduction | The key questions are appropriate and explicitly stated | No action required. |
| **Peer Reviewer #4** | Introduction | It would be helpful to provide an explicit rationale in the introduction for why cancer risk and literacy level were examined as effect modifiers. Why might effectiveness or receptivity vary by these characteristics? For instance, is the assumption that higher risk subgroups are more likely to attend to the issue/engage in the hard work of deliberating over a decision? Likewise for low literacy subgroups - is the assumption that they would be less likely to understand or attend to the information communicated? | We thank the reviewer for the comment. We have now added in the Introduction: "For example, research suggests that patients' baseline understanding of issues in cancer screening may affect whether they ultimately made an informed and considered choice,[ref] and that patients' perception of their own risk was an important predictor of cancer screening uptake.[ref]" |
| **Peer Reviewer #5** | Introduction | The introduction could build a stronger argument for key question #1 - why is it important to consider risk groups and why should aid perform differently. Is the underlying hypothesis that decisions where the stakes are higher (e.g., early treatment) are different than prevention decisions for example, and therefore the aids may be more/less effective? | We thank the reviewer for the comment. We have now added in the Introduction: "Previous research has indicated that understanding the content and purpose of the decision aid affected the participant's potential of reaching an informed choice about screening, while perceived susceptibility was one of the constructs that were observed to be an important predictor of screening uptake." |
| **Peer Reviewer #6** | Introduction | short but provided enough background. | No action required. |
| **Public Reviewer #1 (Oncology Nursing Society)** | Introduction | page 9, line 41, suggest change "..particularly for decisions related to early cancers (decisions related to screening, preventive treatment, or treatment..." to "...particularly for decisions related to screening prevention or treatment of..." | We made this change using the wording suggested by the reviewer. |
| **TEP Reviewer #1** | Methods | Very complex set up studies that were summarized well. | No action required. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #2** | Methods | Methods are appropriate, including and exclusion criteria make sense, definitions used are appropriate. Statistics appear correct although I am not a statistician. | No action required. |
| **TEP Reviewer #2** | Methods | page 11, line 16, change "...favorable prognosis (typically locally not advanced)..." to "...favorable prognosis (typically local disease only, not advanced)..." | We adopted the reviewer's wording. |
| **TEP Reviewer #3** | Methods | Exclusion criteria: inclusion of other than the classical epidemiological rigorous experimental designs such as RCTs. Perhaps something to reflect upon in the discussion. | Now added to the Discussion: "Decision aids are complex interventions, and their successful integration and continued use in routine care depends on many factors, including patient and provider acceptance, system infrastructure, fit with other processes, and other factors only peripherally related to the patient-provider dyad. Thus, implementation of decision aids interventions in routine practice requires consideration of many additional factors.<br>Although we looked for studies of the effectiveness of interventions to providers for promoting shared decisionmaking through decision aids, we found limited evidence. A more general treatment of shared decisionmaking promotion interventions did not draw strong conclusions" |
| **TEP Reviewer #3** | Methods | Search strategies look fine, and are efficiently based on earlier work from others. | No action required. |
| **TEP Reviewer #3** | Methods | Outcomes: page 11: Not a very important issue but I just wonder if the DCS does not fit more to the category of measurements of decisional quality and cognition. | We agree that DCS could fit in either category. For simplicity of exposition we keep it in our originally assigned category. |
| **TEP Reviewer #3** | Methods | Page 13: data extraction: What do you mean with explicit versus implicit elicitation of values? Can you give an example or be more clear on the difference as you defined it. | Now clarified. |
| **TEP Reviewer #3** | Methods | Page 13: data extraction: What do you mean with description of problem and options? There is a world behind this (risk communication science, framing risks in various ways (Gigerenzer, Kahneman). | To avoid confusion, the term "clear" was omitted. |
| **TEP Reviewer #3** | Methods | Page 13: data extraction: Other attributes whether it was based on theory, how did you define this? Do you consider use of IPDAS standard as 'based on theory' | The extracted information was based on reporting at the individual study level and this clarified now. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #3** | Methods | Page 13: data extraction: 'We imputed missing standard deviations as the median standard deviation in less than 8% of arms." Why did you choose to impute missing data? I'm not sure if this is important for the quality of the review. (I have no experience with imputation myself) | We thank the reviewer. We chose to impute missing data based on the previous extensive experience of the group with good imputation performance results. |
| **TEP Reviewer #3** | Methods | Should this Dutch study not be included?<br>Molenaar S, Sprangers MA, Rutgers EJ, Luiten EJ, Mulder J, Bossuyt PM, et al. Decision support for patients with early-stage breast cancer: effects of an interactive breast cancer CDROM on treatment decision, satisfaction, and quality of life. J Clin Oncology 2001;19(6):1676-87.<br>the study has 2 papers;<br>Molenaar S, Oort F, Sprangers M, Rutgers E, Luiten E, Mulder J, et al. Predictors of patients' choices for breast-conserving therapy or mastectomy: a prospective study. Br J Cancer 2004;90(11):2123-30. | Both studies were considered during the full-text screening process. They were excluded for KQ1 as non-RCT papers. |
| **TEP Reviewer #4** | Methods | P11. Outcomes: later in the manuscript you discuss informed choice or congruence (see pg 22 of results) – but do not describe here as one of the outcomes in decisional quality. Would be good to clarify this outcome and how it is defined. | We added: "Outcomes related to measurements of decisional quality and cognition included differences in knowledge scores (about the condition, options, or expected outcomes as defined in each study); number of people making informed choices (people who have adequate knowledge and make a choice); congruence between actual choices and patient values; and number of people with accurate perception of cancer risk." |
| **TEP Reviewer #4** | Methods | Table 1: p 12 and p 23 For the DCS – it is often reported as a threshold or % age who score below 25 (so not a minimally important difference that we are looking for, but rather % age who are below the threshold). This may change some language around Table 1 and other places this is reported. I do not believe that there is consensus on this reporting but would be good to mention. | Table 1 edited according to the reviewer's comment. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #4** | Methods | Study identification p 13: More details on the screening and review would be helpful. For example, how did you determine that all reviewers applied criteria in the same way? Did you calculate a kappa for the second round of 100 abstracts reviewed? How many reviewers were there for the abstracts? Similarly more details for the full text abstraction would be helpful. For example, how were discrepancies between the two reviewers for full-text articles captured and resolved? | We now clarify: "To ensure consistency, all five reviewers performed a calibration exercise and screened the first 200 citations, in two rounds of 100 citations each, using broad inclusion criteria. Disagreements were discussed and analyzed to clarify screening criteria. Once it was deemed that all reviewers were applying the criteria in the same way, we continued with single screening of the remaining abstracts." and "All included papers were assessed for eligibility by two reviewers. Conflicts and questions were resolved by discussion with a third reviewer." |
| **TEP Reviewer #4** | Methods | Data extraction p 13: Given that information on the decision aid (especially content and theoretical framework) may be in other published reports and not in the RCT manuscript did you use other methods to gather that information? If so what did you do? If not, why not? | Now added: "Information on the characteristics of the decision aids and numerical information was extracted or cross-checked at least twice. If that information was not in the paper, we attempted to access the original decision aid or other studies of that decision aid for this information. Only in cases where this was not possible was any information on DA characteristics listed as not reported." |
| **TEP Reviewer #4** | Methods | Data extraction p 13 For the delivery format – is it possible for decision aids to have multiple components and if so how were the data categorized (e.g. we have decision aids that include DVD and booklet – were those counted as both?). Also if the software and/or web site had audio visual was that considered different just because it was viewed on computer screen as opposed to television screen? Do you think the lack of difference in format is because there was double counting (if there was) or because the fundamental categories were not really different? For example, would linear versus interactive make more of a difference? | We now edited the report as follows: "(3) the delivery format or formats (printed, audio or video material not on a computer, computer software, Web site, in-person delivery with a person providing logistical help, use of support groups or patient navigators, decision board/option grid)." |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #4** | Methods | Seems that the search did not find articles for the second key question. Was this because of how narrowly defined the search was? Seems that other reviews of implementation have found many other articles (including ones in cancer). Was this defining shared decision making as decision aids only? Did it include training in communication skills for SDM? | Yes, the limited yield for KQ2 was due to a narrowly defined search looking for decision-aid related interventions. Our rationale for that choice, also supported by strong and explicit TEP and consultant advice, was serving the equilibrium between the risk of scope creep and lack of generalizability (already addressed in the Discussion). We now added: "For the second Key Question, we included comparative studies informing on the effectiveness of interventions for promoting shared decisionmaking to providers caring for the populations discussed for the first Key Question, specifically provider-targeted interventions to increase shared decision making with the use or increased use of a decision aid." |
| **Peer Reviewer #1** | Methods | Apparently, one search strategy was performed to answer two key questions. This can be a fatal flaw to answer either of these questions, because it is doubtful that all relevant trials were found for both questions. | Added under study identification: "We searched MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials (CENTRAL), PsycINFO, and the Cumulative Index to Nursing and Allied Health (CINAHL) from inception to October, 2013, using two separate strategies, one for each key question, which were based on previous Cochrane reviews.[4 10] Both strategies are reported in Appendix A." |
| **Peer Reviewer #1** | Methods | The authors apparently did not use the PRISMA checklist as an internationally accepted standard to ensure a proper conduct and description of the systematic review. | At the end of the first paragraph of the methods section, we added: "In reporting this systematic review, the authors followed the PRISMA guidelines. |
| **Peer Reviewer #1** | Methods | Outcomes were not predefined, but "identified prospectively", which introduces the risk of publication bias (trials may only report on positive results). | We now clarify that outcomes were pre-specified in the protocol. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #1** | Methods | A motivation is lacking why "Quantitative analyses were run for outcomes reported in at least 10 trials overall and in at least 2 trials in each population group". | We now clarify in the report that “Our main analyses used hierarchical (random effects) regression models adjusted for population group (average risk, high risk, early cancer) and additional intervention characteristics. These models can be difficult to fit with few studies. Thus we run analyses in outcomes with at least 10 trials overall and with at least 2 trials in each population group” |
| **Peer Reviewer #1** | Methods | The possibility of, or conditions for, performing meta-analyses is not discussed. | A meta-analysis was done. Please see the description of the meta-analysis in the reply to the previous comment. |
| **Peer Reviewer #1** | Methods | It is unclear whether subgroup analyses were planned in advance (which should be done) or performed post-hoc (which should not). | We identified factors in advance and this is now clarified. |
| **Peer Reviewer #2** | Methods | Methods are all well justified, logical, appropriate and well described. | No action required. |
| **Peer Reviewer #3** | Methods | The inclusion and exclusion criteria are justifiable | No action required. |
| **Peer Reviewer #3** | Methods | The search strategies are explicitly stated and logical. | No action required. |
| **Peer Reviewer #3** | Methods | The definitions for the outcome measures are appropriate | No action required. |
| **Peer Reviewer #3** | Methods | The statistical methods used are appropriate | No action required. |
| **Peer Reviewer #4** | Methods | The inclusion and exclusion criteria are justifiable and appropriate. The search strategies are explicitly stated in the appendix and logical. The statistical methods are rigorous. | No action required. |
| **Peer Reviewer #4** | Methods | The review summarizes a number of outcomes for which the specific definitions vary considerably across studies. The authors have done a nice job of making sense of this heterogeneity. | No action required. |
| **Peer Reviewer #4** | Methods | It would be helpful on page 11 to define "accurate risk perception" more explicitly. Do you mean perception of PERSONAL risk for cancer, death, benefits and harms associated with different option; perception of AVERAGE risk of cancer, death, benefits, harms; or either/both of these? | Changed to “accurate perception of either their personal cancer risk given the available choices or their average baseline risk.” |
| **Peer Reviewer #4** | Methods | The examination of effects across risk groups (average, high, early cancer - page 21) conflates effects due to risk status with effects due to the nature of the decision being made. Decision aids for people without cancer are, as the authors mention, primarily focused on whether or how to be screened. Decisions for early cancer patients are primarily focused on what type of treatment to get. I am not sure how you can separate the effect of risk status from the effect of type of decision being made, or whether it is useful to try to do so. | We acknowledge the reviewer’s point; yet we would argue that in our review we observed no differences in overall vs. stratified analyses. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #4** | Methods | Many decision aids are tailored to specific risk and literacy levels (or designed to be accessible to low literate populations). It would be helpful to clarify how such tailoring was accounted for in your analysis and of interpretation of the findings regarding risk and literacy level as effect modifiers. | Now added: "We examined effect modification for each population group (screening, high risk, early cancer), and for delivery formats, content, other attributes of the decision aid (whether it was interactive, tailored to target population, such as low literacy, or used by consumer and provider together or by the consumer only), and design items" |
| **Peer Reviewer #5** | Methods | The methods seem reasonable and well presented. | No action required. |
| **Peer Reviewer #6** | Methods | Selection of outcomes: the explanation of why hard outcomes are not used is not very clear and may be better clarified. | We clarify that: "Almost by definition, for most situations for which patient decision aids are proposed, the likelihood of mortality or other hard clinical outcomes across the compared options is either known to be similar or is substantially uncertain. Because there is no single optimal choice, hard clinical outcomes are probably not particularly relevant for measuring the effectiveness of decision-aid-based interventions. Intermediate health outcomes, such as quality of life, anxiety, depression, or decisional regret, are more relevant measures of the effects of decision-aid-based interventions." |
| **Peer Reviewer #6** | Methods | For continuous outcomes, did the included papers measure the outcomes at both baseline and after treatment, or only after treatment? Were the differences of follow-up scores used, or differences in change scores used? Clarify the measure used in the analysis. | Baseline measurements were reported in all studies. Follow-up data were reported either as more often follow-up scores or differences in scores used. The endorsed analysis (hierarchical random effects meta-regression analysis) could incorporate both estimates. |
| **Peer Reviewer #6** | Methods | Statistical model - "Associations between the outcomes in each arm" ?? Do you mean differences between arms? State explicitly that a Bayesian model was used and some criteria to choose priors in the primary analyses. | We thank the reviewer for his comment. Due to the considerable length of the review and lack of space, we provide the detailed analysis plan in the Appendix. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #6** | Methods | Sensitivity analysis: Need more clarification on the purpose of the sensitivity analyses related to the Cochrane review. | We now write: "The recent update of the Cochrane review (current as of 2012) included a subset of the trials identified in the current report (see Appendix C for a description of the discrepancy which is mainly because of our including more recent literature). To facilitate comparisons with the conclusions of the Cochrane review, which used different analyses, we repeated all analyses for subset of trials included in the Cochrane review." |
| **TEP Reviewer #1** | Results | The amount of detail is appropriate: neither too little nor too much | No action required. |
| **TEP Reviewer #2** | Results | The results are quite detailed by necessity. I think they are well organized by section and topic which improves greatly the flow of the data. | No action required. |
| **TEP Reviewer #2** | Results | I would like to see some additional explanation provided about the study discussed on page 27, lines 26-30, reference # 83. In this study, use of decision aids (DA) reduced screening participation. I think this outcome deserves more explanation as the DA would appear to have moved patients away from what might be seen as the preferred choice. It is not clear how this contributes to the overall conclusions. | We quote here from the discussion of the paper: "Although the decision aid did not make people more worried about developing bowel cancer, it did make them feel less positive about screening, and reduced uptake of the screening test by 16% (75% in the control group v 59% in the decision aid groups). It seems that this may have resulted from increasing their knowledge about the low personal benefit of screening."<br>We further clarify in the report: "The largest study with statistically significant results for actual choice[83] was done in 572 people in Australia for decisions related to colon cancer screening, and found both increased knowledge and lower rates of screening participation in the decision aid group compared to usual care. The authors attribute this to increased "knowledge about the low personal benefit of screening." |
| **TEP Reviewer #3** | Results | The results section is well-written and clear. The figures are meaningful and interesting, e.g. figure 4. | No action required. |
| **TEP Reviewer #3** | Results | Page 15: 9 studies were cluster RCTs and 38 were multi-center trials. Why this division? Cluster RCTs can be multi-center trials | Now rephrased accordingly. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #4** | Results | P. 23 Cochrane systematic review of decision aids has subgruop analysis on PSA testing trials that did show significant impact of decision aids on uptake of testing, yet here the authors did not replicate that or add to that. Do the authors feel that the Cochrane result is not valid or replicable? Were those studies not included here for some reason? Please comment on the differing results on uptake. | We did not undertake a subgroup analysis specific to PSA testing. Our rationale behind that choice (made among the very large number of potential subgroup analyses) is based on the notion that decision aids are not intended to promote any particular decision and have no impact on the overall distribution in the population. Action and intention outcomes are really of secondary interest in evaluating decision aids unless the measure examined is whether the individual carried out the decision they made. |
| **TEP Reviewer #4** | Results | Tables 3, 4, 5 it would be helpful to include # studies, # participants for the analyses presented in the table if possible. | Edited as suggested |
| **Peer Reviewer #1** | Results | In their Abstract, the authors state: `There were no large differences between using and not using DAs in decisional conflict (DCS) (27 trials, 7,820 participants, weighted mean difference, WMD = -0.22, 95% CrI: -0.38, -0.05)`. `No large` implies an opinion, not a fact. The fact seems to be that there actually is a significant difference. | We thank the reviewer for his comment. Our intent is to summarize evidence and contextualize findings. This amounts to describing our opinions about several aspects, including 1) the risk of bias of each study and of the evidence base; 2) the importance of the findings, including strength of evidence; 3) the generalizability of the findings; 4) the implications for future research. The methods extensively describe how we justify our interpretations and the processes we followed are transparent. |
| **Peer Reviewer #1** | Results | A summary of the risk of bias of the trials included (as detailed in appendix G, and as given for the trials referring to Key question 2) would help appreciate the validity of the available evidence (rather than the strength per outcome as in Table 6). | We think that the RoB, as well as the detailed SOE table belongs to the appendix. The condensed SOE table in the report (Table 6) serves as an overview of the conclusions we draw from this work; we opt to keep it in the Discussion. Secondarily, AHRQ reports typically have such a table in the beginning of the discussion section, and it is desirable to conform to the program's usual expository practice. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #1** | Results | It is unclear what the interventions in the control arms were, if any. | This is clearly described in the Methods Section. The synthesis approach is more elaborate that e.g., Cochrane meta-analysis (that is univariate-response mean-only normal-normal models, fit with 2-step estimation). While we believe the analysis to be appropriate for the problem at hand, we agree with the reviewer's implied comment that it may challenge readers without quantitative background. |
| **Peer Reviewer #1** | Results | The authors have not made, or described, an attempt to extract and interpret the results themselves, but have merely copied the results as given in the papers. | Please refer to the Results section for a description of qualitative analyses for all outcomes and quantitative analyses for outcomes with enough information. |
| **Peer Reviewer #1** | Results | I´m not sure how valid and precise the evolution analysis over time actually is. | The description of the temporal evolution of characteristics of the format and content of the interventions is a description, not an analysis. The information was doubly extracted, and contrasted with Cochrane Tables, when they were available. |
| **Peer Reviewer #1** | Results | Effect modification analyses for several factors seemed frequently underpowered due to small subgroups and therefore does not allow a conclusion about whether effect modification may or may not be present. This should be mentioned in the Limitations section. | We added text in the discussion to clarify the point raised by the reviewer.<br>Results on effect modification have wider credible intervals for knowledge, and relatively narrow credible intervals for DCS and anxiety. |
| **Peer Reviewer #2** | Results | The results are clearly described, present the right amount of information, and are based on all important studies. However, I found the markers at the 100% level in Figures 2 and 3 confusing and do not understand why they are there. | Figures edited to address the reviewer's comment (The figure legend elaborates on the axis meaning.) |
| **Peer Reviewer #2** | Results | There is also an apparent typo at the top of page 31: were the California primary care physicians really in waiting areas? | Corrected to: "The second study cluster randomized 120 California primary care physicians in five clinics to …" |
| **Peer Reviewer #3** | Results | The amount of detail presented in the results section is appropriate but forest plot of the meta-analyses would have been appreciated. | The analyses cannot be depicted in a forest plot; they are results of hierarchical regression models.<br>The challenge is that studies have different characteristics and thus the "forest" plots should be one of conditional means for various factors. |
| **Peer Reviewer #3** | Results | The characteristics of the studies are clearly described? | No action required. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #3** | Results | The key messages are explicit and applicable | No action required. |
| **Peer Reviewer #3** | Results | The figures, tables and appendices are adequate and descriptive but again forest plot of the meta-analyses would have been appreciated. | The analyses cannot be depicted in a forest plot; they are results of hierarchical regression models.<br>The challenge is that studies have different characteristics and thus the “forest” plots should be one of conditional means for various factors. |
| **Peer Reviewer #3** | Results | The investigators did not overlook any studies that ought to have been included or conversely they did not include studies that ought to have been excluded. | No action required. |
| **Peer Reviewer #4** | Results | The amount of detail presented in the results is appropriate and the characteristics of the studies are clearly described. The key messages are explicit and applicable. | No action required. |
| **Peer Reviewer #4** | Results | Figure 4 - I think it would help with the digestion of this table to add labels for the horizontal lines separating trials for average, high risk, and early cancer populations. It might also be helpful to sort the studies within each section of the figure group by author’s last name and publication date. It is not clear how they are currently sorted. | Edited as proposed. |
| **Peer Reviewer #4** | Results | Page 23 - Accurate perception of fatality risk: The choice of the term "fatality" in this context is unusual. "Mortality" would be more normative and is the term used in most decision aids. | Change to : “accuracy of perception of mortality risks . . .” |
| **Peer Reviewer #4** | Results | Page 23 - last paragraph - typo: Insert a period after the word "important". | Done. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #4** | Results | Pages 26 - 27 - Actual or intended choices: I think it would be helpful to use a more neutral framing in discussing the association between decision aid exposure and decision action and intention. Decision aids are not intended to promote any particular decision. Therefore, use of "effectiveness" language to describe these outcomes seems awkward to me. It is possible that a decision aid modifies the intentions of many individuals but has no impact on the overall distribution in the population. It is also possible that a decision aid could lead to more informed, satisfactory, participatory decisions without any change in the individual or population distribution of choices selected. It is not necessarily a failure of a decision aid if exposure to it has no impact on the pre-post distribution of preferences. A few words to make the framing more neutral would be beneficial here. Action and intention outcomes are really of secondary interest in evaluating decision aids unless the measure examined is whether the individual carried out the decision they made. It does not appear that any of the studies reviewed examined this type of outcome. | We rephrased using the reviewer's suggested wording: "Overall, 48 trials examined the ways decision aids impacted actual or intended choices for the decisional problems at hand." |
| **Peer Reviewer #4** | Results | In the results section sometimes specific odds ratios and precision estimates are provided and sometimes they are not. For instance, Page 30 - last paragraph - the authors present the odds ratio for last association referenced in the sentence but not the first. It would be helpful if the authors were more consistent throughout the results in terms of details provided about estimates. | We acknowledge the reviewer's point. This has to do with what is and is not reported in the individual studies that are assessed qualitatively. Nevertheless, for the sake of consistency, the provided estimates were deleted. |
| **Peer Reviewer #5** | Results | A point about decisional conflict and the decisional conflict scale: the initial thinking about this construct and measure was that decision aids should reduce decisional conflict. Somewhere along the way people started talking about how aids might actually increase decisional conflict but they offered no evidence to support these claims. While it is fine to say that aids do not increase decisional conflict, in my opinion the issue is whether or not the aids reduce decisional conflict. | We acknowledge the reviewer's comment and now provide a balanced view on the observed DCS change direction. |
| **Peer Reviewer #5** | Results | This point relates to a second issue. If I recall, the Cochrane review shows that aids lead to reduced decisional conflict for the Informed and Values subscales primarily. I suggest considering looking at these subscales if possible in the review and reconsidering the findings if they are more in line with the Cochrane review. | Under extensive discussion and TEP support, we chose in our analysis plan not to use subscale-specific analyses. Doing the analyses correctly requires multivariate meta-analysis, and us getting a hold of the correlations between the scale responses. We have asked the scale developers for this information but they have not studied it. |
| **Peer Reviewer #5** | Results | I'm having trouble understanding the circles in Figures 2 and 3 despite the description. This needs more work | As suggested, we edited the figure legends and enhanced image resolution. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #5** | Results | In Figure 3, I suggest defining "content" for the reader. | Now elaborated. |
| **Peer Reviewer #6** | Results | Figures 2 and 3 are helpful. | No action required. |
| **Peer Reviewer #6** | Results | Page 23 and 27 Define what are “analyzable data” -- it helps to evaluate selection bias or the absence of selection bias. | Syntax corrected. |
| **Peer Reviewer #6** | Results | Tables 3, 4 and 5 – provide the numbers of studies and subjects for each estimate. Also it would be very useful for the readers to see the forest plots of the three meta-analyzed outcomes, stratified by the three populations, if possible, to have a better idea about statistical heterogeneity and how consistent the results are across studies. It also helps to add the scale (for measuring the outcome) of each study in the plots. | Edited as suggested. |
| **Peer Reviewer #6** | Results | Similarly, it also helps to provide some representative numbers in the text, or numbers in a plot (without quantitative synthesis) for some other outcomes to provide some idea about the magnitude and heterogeneity of the differences. It is not easy to do given the diversity in outcomes and reporting, but it is also not easy to evaluate the bottom line for the results of other outcomes. | As the reviewer notes, this is difficult for all the outcomes we are summarizing.<br>The difficulty is compounded by the short format of this review.<br>We have added some more information in selected paragraphs, but it is not possible to add much more information without turning the text into a catalogue of study results.<br>Instead we reference the link to the SRDR site in several places including the methods, beginning of results, and in the paragraphs for knowledge, DCS and anxiety. |
| **Peer Reviewer #6** | Results | Anxiety: Multiple instruments were used among studies and WMD was used in the analysis? What are the scales of the instruments? The combined WMDs are very small – are they actually SMD? | We have amended the text for clarity. We report results from a synthesis of studies using STAI, in a 20-80 range. For this outcome, a WMD is appropriate.<br>In sensitivity analyses we pooled across all instruments using SMDs. The results were very similar with the WMD analyses (practically 0 difference) and are not reported in detail. |
| **Peer Reviewer #6** | Results | What instruments were used for quality of life? | Instruments now provided. |
| **TEP Reviewer #1** | Discussion/ Conclusion | The author did an excellent job summarizing a very diverse set of studies. | No action required. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #2** | Discussion/ Conclusion | In the conclusion, the authors do a good job of discussing the limitations, not just of this project but more importantly of the complexity of evaluating DA given the wide variety of DA used. They also provide general discussion regarding how the development of DA might best progress in the future to overcome these limitations. The author's should add comments regarding specific recommendations for future research, particularly since, having done this exhaustive review, they are well positioned to suggest a best choice path forward to develop, evaluate and implement these tools in the future. | Edited as proposed. |
| **TEP Reviewer #3** | Discussion/ Conclusion | Discussion is of high quality. | No action required. |
| **TEP Reviewer #3** | Discussion/ Conclusion | Page 33: issues on vulnerable population; I'm surprised that coaching by or conjunction with a provider did not have a positive mediating effect on patients with low health literacy. Can you reflect on that? | We have added a comment in the 3rd paragraph of the discussion. |
| **TEP Reviewer #3** | Discussion/ Conclusion | Page 33: you recommend evaluating decision aids for other types of cancer. Are you sure that for each different decision we have to prove effectiveness of decision aids for patients again and again? Do you not think that we evolve towards regarding patient decision aids as 'natural' phenomena, comparable to clinical practice guidelines for professionals? | We thank the reviewer for this insightful comment. We now added in the Discussion: "Thus one might consider the notion of investing in a generic platform for developing and delivering decision aids. The platform could allow for modular expansion of the decision aid content (e.g., to add stories of other people facing a similar problem, or a value clarification exercise) on include web-based ones. It would facilitate development of decision aids in other diseases by removing the need to obtain know how in the technical aspects of the development; translation to other languages; and keeping them current. Clarity in language and accuracy of information according to the IPDAS development process is important and can be achieved perpetually. Moreover, evaluation of performance in real life setting (not necessarily an RCT) is important in itself. Data coming from surveillance/post-marketing studies could help refine the decision aid to the extent possible." |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **TEP Reviewer #4** | Discussion/ Conclusion | Discussion Table 6 p 32 – the Cochrane shows that decision aids decrease decisional conflict. The analyses here seem to downplay or not confirm that result for cancer trials. Would be good for authors to comment on this difference in results. | In our analysis, we also see an effect of a similar magnitude, despite the fact that we perform different analyses. We characterized the effect as a small effect. We have amended the text to make this clear. |
| **Peer Reviewer #1** | Discussion/ Conclusion | The methodological weaknesses in this review should either be corrected or discussed as `Limitations` | I think we have corrected all of these. |
| **Peer Reviewer #1** | Discussion/ Conclusion | The authors have mentioned the likelihood of publication bias, but have not tried to minimize it by predefining the desired outcomes in their review. | I think the fix above should address this. |
| **Peer Reviewer #2** | Discussion/ Conclusion | Yes to all questions: the discussion and conclusion sections are appropriate, grounded in the evidence and constructive. | No action required. |
| **Peer Reviewer #3** | Discussion/ Conclusion | The implications of the major findings are clearly stated. | No action required. |
| **Peer Reviewer #3** | Discussion/ Conclusion | The limitations of the review/studies is described very adequately. | No action required. |
| **Peer Reviewer #3** | Discussion/ Conclusion | Excellent appraisal of the limits of the evidence-base. | No action required. |
| **Peer Reviewer #3** | Discussion/ Conclusion | The investigators did not omit any important literature. | No action required. |
| **Peer Reviewer #3** | Discussion/ Conclusion | There is no specific "future research" section but the authors mention the area with lack of knowledge. | Edited as proposed. |
| **Peer Reviewer #4** | Discussion/ Conclusion | The implications of the major findings are clearly and thoughtfully stated. The limitations are described adequately. | No action required. |
| **Peer Reviewer #4** | Discussion/ Conclusion | I did not find a separate section on future research. There are a few recommendations related to standardized reporting and measurement sprinkled throughout the discussion, but it might be helpful to add a separate section on recommendations for future research. | Edited as proposed. |
| **Peer Reviewer #4** | Discussion/ Conclusion | Page 34, last paragraph, 4th sentence - Typo - I assume "dose" should be "does". | Fixed to: “When independent replications do not exist . . .” |
| **Peer Reviewer #4** | Discussion/ Conclusion | Conclusions - It would be helpful to integrate the information on strength of evidence into the conclusions. | Now added: “We found that decision aids increase knowledge without adverse impact on decisional conflict, or anxiety with moderate to high strength of evidence.” |
| **Peer Reviewer #5** | Discussion/ Conclusion | The discussion is thoughtful and appropriate. I believe the authors are in error about one point. Some specific aids have been evaluated in more than a single study (e.g., the FIDMD aids for CRC and prostate cancer). FIMDM allowed other research to use/evaluate their tools, so the point about the same teams evaluating their own tools seems in error | We thank the reviewer for this comment. The statement, now clarified, pertains to the included RCTs and does not stand for other study designs evaluating the feasibility of implementation. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #6** | Discussion/ Conclusion | Other than no validation study, what are the potential implications of 54 decision aids that were used? Or it is a matter that there are just many decisions aids around? For example, based on Table 2, for the population of average risk of cancer, 39 aids estimate the generic risk probabilities but 22 studies were for prostate cancer. So it seems like that many aids may serve the same or similar purpose? Any need for future work to see which aids work better? | We thank the reviewer for his valid point. We now added in the Discussion: "Thus one might consider the notion of investing in a generic platform for developing and delivering decision aids. The platform could allow for modular expansion of the decision aid content (e.g., to add stories of other people facing a similar problem, or a value clarification exercise) on include web-based ones. It would facilitate development of decision aids in other diseases by removing the need to obtain know how in the technical aspects of the development; translation to other languages; and keeping them current. Clarity in language and accuracy of information according to the IPDAS development process is important and can be achieved perpetually. Moreover, evaluation of performance in real life setting (not necessarily an RCT) is important in itself. Data coming from surveillance/post-marketing studies could help refine the decision aid to the extent possible." |
| **Peer Reviewer #6** | Discussion/ Conclusion | For the conclusion of "no difference in effectiveness of decision aids by examined characteristics": Overlap of 95% CIs and "explained (heterogeneity in ??) effectiveness of decision aids (page 35, lines 39-41) are two different things. | Edited as follows: "A contribution of our systematic review is that it explicitly examined differences in the effectiveness of decision aids by isolating attributes of their delivery format, content, and other factors, and found that the currently accumulated randomized evidence does not support an association between isolated attributes and decision aid effectiveness." |
| **Peer Reviewer #6** | Discussion/ Conclusion | For the conclusion of "no difference in effectiveness of decision aids by examined characteristics": Meta-regression analyses are typically inadequate powered – is this a potential reason for the lack of difference in effectiveness by the examined characteristics? | We agree. Edited as follows: "A contribution of our systematic review is that it explicitly examined differences in the effectiveness of decision aids by isolating attributes of their delivery format, content, and other factors, and found that the currently accumulated randomized evidence does not support an association between isolated attributes and decision aid effectiveness." |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #6** | Discussion/ Conclusion | For the conclusion of "no difference in effectiveness of decision aids by examined characteristics": Differences in effectiveness of decision aids by characteristics are only systematically examined in three outcomes, but the conclusion was implied for the general effectiveness of decision aids. | We believe that the conclusion does not imply this for other outcomes. For example, see the SOE Table (summary in Table 6, detailed in the Appendix). |
| **TEP Reviewer #1** | Clarity and Usability | The report is well structured and organized. | No action required. |
| **TEP Reviewer #2** | Clarity and Usability | I think this report represents an excellent review of a complex topic. It is well organized and complete. | No action required. |
| **TEP Reviewer #3** | Clarity and Usability | What do you mean with "with factor and without factor' in the tables? | Syntax corrected. |
| **TEP Reviewer #3** | Clarity and Usability | Not all other attributes in table 3 are mentioned in the methods section. | Edited as proposed. |
| **TEP Reviewer #3** | Clarity and Usability | Ref 29 incomplete, Ref 76 use of capitals | Corrected. |
| **TEP Reviewer #3** | Clarity and Usability | Why not a crude data table listing all studies? | We acknowledge the reviewer's point. The complete crude data are available in SRDR. In addition, we will provide a supplementary crude data file for alternative access. |
| **TEP Reviewer #3** | Clarity and Usability | You use 'decision-making' and 'decisionmaking' in the manuscript, the latter looks strange for me. But must be due to my non-native english speaking status. Same for credible intervals while I'm used to confidence intervals. | Corrected.<br>We refer to (Bayesian analysis) credible intervals. |
| **TEP Reviewer #3** | Clarity and Usability | Page 21: An SMD of 0.20-0.30, typing error? | Changed to "0.20 to 0.30". |
| **TEP Reviewer #3** | Clarity and Usability | Page 23, last paragraph: you lack bullet point | Fixed. |
| **TEP Reviewer #3** | Clarity and Usability | Page 31, first paragraph of discussion: should the word 'often' not be 'which'? | Corrected: "The assessed decision aids showed considerable heterogeneity in terms of format, content, context and theoretical background, which often made synthesis a challenge." |
| **TEP Reviewer #4** | Clarity and Usability | The report would benefit from some clarifications as described above. Generally, the report is well written and clearly states the key results. Given the lack of studies that they found for the second questions (and potential missing studies) wonder whether it makes sense to include? | We acknowledge the reviewer's point. Yet, KQ2 was *a priori* defined and we chose to show the observed lack of evidence. |
| **TEP Reviewer #4** | Clarity and Usability | The conclusions can inform policy as they show little evidence in many areas and the great need for more systematic investigation. | No action required. |

| Commentator & Affiliation | Section | Comment | Response |
|---|---|---|---|
| **Peer Reviewer #1** | Clarity and Usability | The report is well organized, but the conclusions may not be valid, given the apparent methodological shortcomings in the review. | We thank the reviewer for the time and effort taken to review our work. In our response, we have addressed in detail all the reviewer's comments, and we anticipate that the clarifications provided resolve any … and show that the suggested alternatives were extensively discussed and deemed suboptimal. |
| **Peer Reviewer #1** | Clarity and Usability | It is unclear what the additional value of this review is over other existing ones. | We now further elaborate: "A contribution of our systematic review is that it explicitly examined differences in the effectiveness of decision aids by isolating attributes of their delivery format, content, and other factors, and found that none is associated with decision aid effectiveness. This suggests that simpler (and less costly to develop and maintain, and easier to use) decision aids may be as effective as more complex ones" |
| **Peer Reviewer #2** | Clarity and Usability | I found the report very clearly written and the reports a good guide for future research in the area. | No action required. |
| **Peer Reviewer #3** | Clarity and Usability | The report is well structured and organized. | No action required. |
| **Peer Reviewer #3** | Clarity and Usability | The main points are clearly presented. | No action required. |
| **Peer Reviewer #3** | Clarity and Usability | The conclusions can be used to inform policy and/or practice decisions. | No action required. |
| **Peer Reviewer #4** | Clarity and Usability | The report is well structured and organized and the main points are clearly presented | No action required. |
| **Peer Reviewer #4** | Clarity and Usability | I think the conclusions would be more actionable if strength of evidence information was integrated into them, and the areas where the evidence is insufficient / more research is needed to inform policy were more clearly defined | We believe that the SOE Table 6 gives that information. Moreover, we now added: "We found that decision aids increase knowledge without adverse impact on decisional conflict, or anxiety with moderate to high strength of evidence." |