Author's response to reviews

Title: Shared decision making for patients with type 2 patients: a randomized trial in primary care

Authors:

Megan E Branda (branda.megan@mayo.edu)
Annie LeBlanc (leblanc.annie@mayo.edu)
Nilay D Shah (shah.nilay@mayo.edu)
Kristina Tiedje (tiedje.kristina@mayo.edu)
Kari Ruud (ruud.kari@mayo.edu)
Holly Van Houten (vanhouten.holly@mayo.edu)
Laurie Pencille (pencille.laurie@mayo.edu)
Marge Kurland (mkurland@olmmed.org)
Barbara Yawn (byawn@olmmed.org)
Victori M Montori (montori.victor@mayo.edu)

Version: 5 Date: 29 June 2013

Author's response to reviews: see over
Rochester, June 28, 2013

BMC Health Services Research
RE: Manuscript No. 1927800258879920
Dear Editors

Enclosed please find our manuscript entitled **Shared decision making for patients with type 2 diabetes. A pilot, clustered randomized trial in primary care** revised extensively in response to the editor and reviewers’ comments.

This original manuscript reports on clinical trial results of the impact of decision aids use with type 2 diabetic patients and primary care clinicians compared to usual care in the discussion of medication management. The results come from patient and clinician reported outcomes, medical records and pharmacy records that were obtained from the randomized trial of decision aids (Diabetes Medication Choice and Statin Choice). This work builds on the body of evidence that evaluates decision aids and the role of shared decision making within the clinical encounter, evaluating the outcomes and experience from both the patient and clinician perspective.

Below is a list of modifications that were made after addressing the reviewers’ comments; these changes were also highlighted in the text for ease of review. We have also provided point-by-point responses to the reviewers’ concerns below.

Sincerely,

Victor M. Montori, MD, MSc
Professor of Medicine
Mayo Clinic
Professor of Clinical Epidemiology and Biostatistics
McMaster University
Changes to manuscript:

1. Page 2, Para 2: Modified text in the Methods section; Before: ‘In random concealed fashion, 10 rural primary care practices enrolled 103 patients with type 2 diabetes and implemented either a decision aid about starting statins or one about choosing antihyperglycemic agents while serving as a control group for the other decision aid.’ After: ‘We cluster-randomized 10 practices in a concealed fashion to implement either a decision aid (DA) about starting statins or one about choosing antihyperglycemic agents. Each practice served as a control group for another practice implementing the other type of DA. From April 2011 to July 2012, 103(DA=53) patients with type 2 diabetes participated in the trial.’

2. Page 2, Para 2: Modified text in the Methods section; Before: We used post-visit patient and clinician surveys to collect decisional quality outcomes (patient knowledge and comfort with decision making) and decision-making process outcomes (patient and clinician satisfaction), and health outcomes (glycemic and lipid control, adherence to therapy) from medical records and pharmacy fill profiles.’ After: ‘We used patient and clinician surveys administered after the clinical encounter to collect decisional outcomes (patient knowledge and comfort with decision making, patient and clinician satisfaction). Medical records provided data on metabolic control. Pharmacy fill profiles provided data for estimating adherence to therapy.’

3. Page 2, Para 3: replace ‘with’ with ‘to’ in the sentence ‘Compared to usual care..’


5. Page 2, Para 3: added p-values to each reported estimated differences

6. Page 2, Para 3: added ‘by their clinicians’ to clarify outcome of ‘engaged by their clinicians in decision making’

7. Page 2, Para 3: replaced ‘without’ with ‘We found no’

8. Page 2, Para 3: Added to sentence ‘, in part due to limited statistical power’.

9. Page 3, Para 1: Replaced conclusion of: Before ‘SDM is challenging yet feasible in rural primary care clinics. Decision aids designed for point-of-care use improve the quality of decision making in patients with type 2 diabetes.’ After: ‘DAs improved decision outcomes without significant effect on clinical outcomes. DAs designed for point-of-care use with type 2 diabetes patients promoted shared decision making in nonacademic and rural primary care practices.


12. Page 4, Para 2: added references to the Center for Medicare and Medicaid services sentence and the National Quality Forum sentence.

13. Page 4, Para 2: added sentence to end of paragraph – ‘These legislative efforts require improvements in the evidence base about the effect of implementing SDM in usual clinical settings.’
14. Page 4, Para 2: removed sentence ‘Thus, this study sought to contribute knowledge about the feasibility and effectiveness of implementation of SDM to improve the value of healthcare for patients with type 2 diabetes.

15. Page 5, Para 1: added paragraph ‘Thus, the objective of this study was to evaluate, in a cluster-randomized practical trial enrolling nonacademic and rural primary care practices and their patients with type 2 diabetes, the impact of patient decision aids vs. usual care on decision making measures, metabolic control and medication adherence. We opted for a cluster-randomized trial, with randomization at the clinic level in order to reduce the risk of contamination, mitigate confounding with clinician communication style, and facilitate the implementation of study procedures and of the decision aid in each practice [12].

16. Page 5, Para 2 under ‘Methods’: replaced ‘rural and suburban’ with ‘nonacademic and rural’

17. Page 5, Para 2: capitalized ‘Southeastern’

18. Page 5, Para 2: replaced ‘which’ with ‘These procedures’ and added ‘these’ to for ‘present these here briefly’.

19. Page 6, Para 1: added ‘care’

20. Page 6, Para 1: added ‘Practices were deemed eligible if they provided primary care for patients with type 2 diabetes.’

21. Page 6, Para 1: added ‘, identified by their clinician,’

22. Page 6, Para 1: added ‘For example, for the diabetes discussion, eligible patients had HbA1c > 7.3, were not using insulin, and were not taking > 2 antihyperglycemic agents at maximum dose. For the discussion about statins, eligible patients should not be using statins and should not have contraindications for taking statins.’


24. Page 8, Para 1: replaced ‘exchanged’ with ‘present’

25. Page 8, Para 2: replaced ‘at’ with ‘after’

26. Page 9, Para 2: added ‘starting’ to sentence

27. Page 9, Para 2: replaced ‘to’ with ‘and spanning up to’

28. Page 9, Para 2: replaced ‘by calculating’ with ‘This estimate of adherence was calculated using’

29. Page 9, Para 3: replaced ‘for’ with ‘to adjust’

30. Page 10, Para 1: replaced ‘Adjusted’ with ‘We report adjusted’ and removed ‘are reported’ from end of the sentence.

31. Page 10, Para 1: replaced ‘adjusting’ with ‘to adjust’

32. Page 11, Para 1: removed reference to unpublished work

33. Page 11, Para 1: replaced ‘their’ with ‘the’

34. Page 11, Para 1: removed ‘from the main analyses’

35. Page 11, Para 1: replaced ‘(their results are presented in appendix B)’ with ‘we present their results in appendix B.’
36. Page 11, Para 3: moved results for satisfaction results to beginning of paragraph from end. To reflect order of results in tables.

37. Page 12, Para 1: replaced ‘patient engagement’ with ‘clinician effort to engage patients in decision making’

38. Page 12, Para 1: moved results for video recordings on fidelity from clinician outcomes to last sentence in ‘Decisional outcomes-patient outcomes’ to reflect order of results in tables.

39. Page 12, Para 1: added ‘in the OPTION score’

40. Page 12, Para 2: replaced ‘this’ with ‘the difference in the frequency of concordance’.

41. Page 13, Para 2: removed ‘their clinician’

42. Page 13, Para 2: added ‘have their clinician put effort into engaging them’

43. Page 13, Para 2: added ‘; these estimates, though, were imprecise due to low statistical power.’

44. Page 13, Para 2: removed ‘The lack of significance found in these outcomes is not interpretable as a true lack of difference due to the study being underpowered. The use of the decision aids was feasible for patients and clinicians in these rural practices. Two qualitative studies of the use of decision aids and conduct of trials in these practices are reported elsewhere. [10, Unpublished Observations Ruud 2012]’

45. Page 13, Para 3: removed sentence: ‘Some of our estimates were imprecise thanks to the small size of the trial.’

46. Page 13, Para 3: added ‘; this reduced the precision with which we were able to estimate the intervention effect on clinical outcomes. We were able to obtain video recordings from 38% of encounters. This limits our ability to use our checklist and to obtain an OPTION score in all encounters thus reducing our confidence in the inferences related to fidelity and clinicians’ efforts to engage patients in decision making, respectively.’

47. Page 14, Para 1: removed reference to unpublished work

48. Page 14, Para 1: added ‘While we did not appreciate important clustering effects, our trial has all the challenges of clustered-randomized trials in that the small number of clusters reduces the likelihood that randomization will achieve balanced prognosis. Also, incomplete recruitment of clinicians and patients within each cluster might introduce selection bias, particularly through incomplete allocation concealment once sites have been allocated but clinicians and patients are still being recruited.’

49. Page 14, Para 2: added ‘, particularly on decisional outcomes,’

50. Page 15, Para 2: replaced ‘important implementation changes are resolved.’ with ‘the decision aids are implemented.’

51. Page 16, Para 1: added ‘and rural’

52. Page 16, Para 1: added ‘and’

53. Page 16, Para 1: removed ‘and have not had favorable impact on adherence or clinical outcomes.’

54. Page 16, Para 1: added ‘This effectiveness has not necessarily translated into favorable impact on clinical measures of effect or on medication adherence. We have seen this not only here, but also in our prior studies and on the recent Cochrane review of decision aids [2]. This might result from two closely related issues. The first issue is that decision
aids that operate during the consultation require attending the consultation, itself a manifestation of adherence. Thus, without invoking trial selection bias, the requirement for our decision aids to operate in the consultation may limit their exposure to non-adherent patients. This limits the opportunity to find effects on medication adherence. The second closely related issue is that shared decision making is appropriate when there is more than one sensible management option. When one option yields superior outcomes with acceptable harms and costs, then most patients will select this option, and the outcomes achieved by the group will reflect the effect of each patient implementing that option. Any variability observed across patients will result from chance, differences in biological response, and treatment adherence. When more than one option is acceptable, however, the effect of different patients opting for different treatments will contribute to the overall variability in outcomes. Therefore, in the absence of a superior treatment choice, the only way shared decision making can improve patient outcomes is through improvements in adherence to the selected treatment option. Thus, there is no a priori reason to believe that shared decision making should improve clinical outcomes beyond its effect on adherence to treatment.

55. Page 17, Para 1: added ‘We have discussed above the fact that the observed fidelity of use might suggest that some effectiveness might not have been realized because of insufficient clinician training. An alternative explanation is that patient-centered care sometimes necessitates deviation from the expected use of the tools in order to accommodate emerging patient issues. Thus, whether this means that more effort to improve optimal use of the tools is necessary warrants further exploration.’

56. Page 17, Para 2: replaced ‘Yet their value as promoters’ with ‘Despite the uncertainty about their impact on clinical outcomes, the value of decision aids as promoters..’

57. Page 17, Para 2: added ‘DAs promote patient-centered practice to the extent that they support both parties in having an evidence-based discussion in which patient participation in deliberation is dynamically and empathically negotiated by the parties. DAs do not guarantee patient-centered care to the extent that the practices, norms, rituals, and policies of the practice may fail to support it [22]. Similarly patient engagement is facilitated by the common ground offered by the decision aid, but it might not happen if the patient is not in a position to participate or feels threatened by such participation [23]. Thus, DAs are tools to promote and facilitate participatory forms of decision making, but much work is needed to increase the likelihood that shared decision making results from their use.’

58. Page 18, Para 1: replaced ‘these results will satisfy’ with ‘the results from this and past trials will satisfy’

59. Page 18, Para 2: replaced ‘suburban and rural nonacademic’ with ‘nonacademic and rural’

60. Page 21, References: Added references 10, 11, 12

61. Page 22, References: added references 21, 22

62. Modified all footnotes to reflect the updated numbering.

63. Table 1a, removed individual %’s from each row and added a sub-column header to reflect statistics shown ‘n(%)’

64. Table 1b added type of clinician (Provider, Resident/Fellow, NP/PA)
65. Table 2a modified title so it now reads ‘Decisional Outcome – Knowledge assessment’
66. Table 2a added column to reflect mean difference with 95% CI along with p-value where appropriate. Added footnote 6 to reflect type of p-value.
67. Table 2b. Modified title so it now reads ‘Decisional Outcome – Satisfaction with knowledge transfer’
68. Table 2b. added p-value for each comparison and added footnote to reflect the statistic used to compute p-value.
69. Table 3. Modified title to now read ‘Decisional Outcome – Decision and Comfort with Decision’
70. Table 3. Removed footnote that read ‘DCS and DCS sub-scales’ and added title to each row that the DCS outcomes were reflected.
71. Table 3. Added p-value and mean difference along with (95% CI) where appropriate.
72. Table 4. Modified title to read ‘Decisional Outcome – Primary Care Clinician Results’
73. Table 4. Added p-value and footnote 3 to reflect the statistic used to calculate p-value.
74. Table 5. Modified title to read ‘Clinical Outcomes – Lab Results and Medication Adherence’
75. Table 5. Added p-values and footnote 5 to reflect test statistic used to calculate p-value
76. Table 5. Modified footnote 4 to reflect model
77. Figure 1. Modified diagram to reflect clustering.

Response to editor comments:
1. There are two Figure 3s. Which is to be included in the paper?
   • Removed incorrect figure.
2. Figure 3 needs to be corrected to accurately reflect the design - it gives the impression that patients were randomized to conditions.
   • Figure has been updated to reflect clustering design.
3. References to Unpublished Observations should be deleted as they are gratuitous. The authors should simply state those observations in their text as their own work.
   • All references to unpublished work has been removed.
4. The text refers to Appendices in two places. Those appendices are not apparent from the included supplements. They should be included and appropriately identified or the references should be deleted.
   • Both appendixes have been added and titled.
5. The manuscript refers to the use of regression analyses but does not present any tabular information reporting those analyses. These tables should be included in the manuscript.
   • Footnotes have been modified to reflect the type of analysis completed.
6. Please respond to or rebut comments/issues by reviewer 2:
   • See below.

Response to reviewer 1 (receipt of review 5/19/2013)
Reviewer 1:
No modifications were needed by reviewer.
Response to reviewer 2
Reviewer 2 general comments:
I have read the revised paper. Some things are clearer. The decision to delete the information on feasibility is an interesting one. I appreciate the inclusion of the articles in press. However, I have not read them in detail for this review, since this is a burden to prospective readers of the manuscript under review. While publishing the same article in multiple places is not appropriate, any article has to stand on its own and necessary context must be provided. I note that while a different article has been published on feasibility, the authors continue to discuss it here, but with inadequate supporting evidence. I will note where this discussion must be deleted to maintain the focus and coherence the authors state they desire (see their cover letter).

In general, there is much detail not reported that makes the article difficult to understand and evaluate. To inform readers and advance the field, accurate and more transparent description is needed. These will not require many more words. Now that the focus has been narrowed, space should not be a problem. Both reviewers have noted this problem. The revised manuscript does not adequately address this problem.

Response to reviewer:
We thank the reviewer for the in-depth review of our paper. We agree with the reviewer on many points and have done our best to address all concerns. We are trying to provide as much detail and information as we can, as demonstrated by our publication of a protocol paper, a report of the main quantitative results (this one), a paper with qualitative results, and a paper about the lessons learned in trying to complete this difficult trial. Furthermore, our research team has contributed to the CONSORT statement and to many publications about appropriate reporting of trials. To this extent we are quite grateful to have limitations in our efforts pointed out so that we can produce the best possible report of our work.

Below is a detail listing of responses to each one of the reviewer’s comments:

<table>
<thead>
<tr>
<th>Reviewer Comment</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>From the &quot;lessons learned&quot; paper and the qualitative paper, my brief read of the abstracts suggests that this trial shows that DAs do not necessarily produce shared decision making and that getting them used at all is challenging. In that case, this report of the results of the trial needs to be more forthcoming.</td>
<td>The 'lessons learned' paper addressed trial implementation issues and not the implementation and effectiveness of the decision aid itself, the subject of this submission. The qualitative paper studied the recalled experience of patients and clinicians and does not report on the effectiveness of the decision aid. Thus, there is no substantive overlap in content across these three reports. The efficacy results actually find the decision aids to be effective in producing shared decision making, with outcomes having different statistical power to find such an effect. We point out where this limitation exists quite explicitly.</td>
</tr>
<tr>
<td>It basically shows no impact in clinical results because DAs did not impact clinical practice.</td>
<td>This randomized trial found statistically significant improvements in patient knowledge scores, OPTION scores (indicating clinician engagement of patients), and in the frequency with which a conversation about the choices took place, all benefiting the intervention arm. We did not find significant improvements in clinical outcomes (lab values, medication adherence), but the trial was not powered to these outcomes. The decision</td>
</tr>
</tbody>
</table>
aids, therefore, produced information transfer, engaged patients in decision making, and improved the likelihood of a conversation about choices; i.e., they produced shared decision making. The impact of this intervention on clinical outcomes remains uncertain. This is clearly stated now throughout the manuscript.

Clinician and practice recruitment was very low. Selection bias is highly likely, but not acknowledged.

For all the benefits of cluster trials, there are three major challenges: number of clinics (clusters), difficulty in balancing the number of participating clinicians and patients across sites and limited ability to avoid selection bias or implement allocation concealment within clusters. Our group acknowledges these limitations in a prior publication of another one of our trials ([http://www.trialsjournal.com/content/10/1/113](http://www.trialsjournal.com/content/10/1/113)) indicating our awareness of this issue. The reviewer correctly points out that we have failed to acknowledge these limitations in this report and we do so now in PAGE 14, PARA 1 as follows: While we did not appreciate important clustering effects, our trial has all the challenges of clustered-randomized trials in that the small number of clusters reduces the likelihood that randomization will achieve balanced prognosis. Also, incomplete recruitment of clinicians and patients within each cluster might introduce selection bias, particularly through incomplete allocation concealment once sites have been allocated but clinicians and patients are still being recruited."

Among those recruited to the study, a substantial fraction of the patients did not want to be video recorded, and there is no measure reported of actual shared decision making or of patients’ perception of their own level of engagement. So we really don’t know very much about SDM from this trial.

We appreciate the opinion of the reviewer in this regard. This is an important limitation of our study from the standpoint of fidelity of the intervention. If any, low fidelity tends to move results toward the null. Yet, our results are positive and consistent with previous trials of the decision aids tested here in which the rate of video recordings was higher. This is now acknowledged in PAGE13 PARA 3 as follows: “We were able to obtain video recordings from 38% of encounters. This limits our ability to use our checklist and to obtain an OPTION score in all encounters thus reducing our confidence in the inferences related to fidelity and clinicians’ efforts to engage patients in decision making, respectively.”

A good patient information brochure might have been as effective, as only knowledge was really impacted. These are terribly important issues to report clearly and candidly in the literature. The manuscript, as presently written, does not do this, and is, by omission misleading and does not fulfill its promise to guide the field. Failed experiments can be very important and transformative to a field. But to do so, they have to “own” the results and reflect candidly on the way forward.

Knowledge gains indicate information shared during the clinical conversation. Patients in the decision aid group were significantly more likely to report that this conversation took place (DA=77% vs. UC = 45%). The level of engagement of patient by the clinician (OPTION) was significantly higher with the decision aid amongst recorded encounters. These significant results took place despite low statistical power and potentially low rates of use of the decision aid as indicated by the reviewer previously. In previous trials, these decision aids were compared to brochures and were found superior to brochures in all measures of knowledge.
transfer and decision making processes (Weymiller et al. Arch Intern Med 2007; Mullan et al. Arch Intern Med 2009). Thus, this concern of the reviewer is not supported by the evidence from this and previous trials.

<table>
<thead>
<tr>
<th>Mandatory Revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>The relevant hypotheses prior to beginning the study must be stated. (It is not good enough to say they are published elsewhere.) On page 12, the authors state that patients were blinded to the hypotheses. So are the readers of this report as presently written. The hypotheses will provide important information for understanding the outcomes of the study. Please state the hypotheses related to both clinical outcomes and quality of decision making.</td>
</tr>
<tr>
<td>The reviewer is absolutely right about this deficit of our report. To address the blinding of the readers to our hypotheses, we have added the following text to the Background section Page 5 Para 2 of our manuscript: “Thus, the objective of this study was to evaluate, in a cluster-randomized practical trial enrolling nonacademic and rural primary care practices and their patients with type 2 diabetes, the impact of patient decision aids vs. usual care on decision making measures, metabolic control and medication adherence.”</td>
</tr>
<tr>
<td>Please delete feasibility throughout the paper, beginning with the abstract, as feasibility is no longer to be addressed here by the authors’ intent. Leaving it in with no supporting information is frustrating to the reader.</td>
</tr>
<tr>
<td>We appreciate this call for a parsimonious report and we have removed mentions of our challenges in implementing the trial or the decision aids during the trial as follows:</td>
</tr>
<tr>
<td>Removed from the Abstract: SDM is challenging yet feasible in rural primary care clinics.</td>
</tr>
<tr>
<td>Removed from the Background: Thus, this study sought to contribute knowledge about the feasibility and effectiveness of implementation of SDM to improve the value of healthcare for patients with type 2 diabetes.</td>
</tr>
<tr>
<td>Removed from the Discussion: The use of the decision aids was feasible for patients and clinicians in these rural practices.</td>
</tr>
<tr>
<td>Removed reference to ‘lessons learned’ report from the Results and Discussion.</td>
</tr>
</tbody>
</table>

<p>| Abstract: |
| Methods needs a trial design statement. |
| We agree and have provided a new paragraph with this description (PARA 2, PAGE 2): “We cluster-randomized 10 practices in a concealed fashion to implement either a decision aid (DA) about starting statins or one about choosing antihyperglycemic agents. Each practice served as a control group for another practice implementing the other type of DA.” |</p>
<table>
<thead>
<tr>
<th>Results must include p values for all results, not just selected ones. Where numbers are given, please indicate what is measured (and what is the tool….ie decisional conflict, option, etc)</th>
<th>We have modified the Results section to now state more clearly and completely the results as follows Page 2 Para 3: “Compared to usual care, patients receiving the DA were more likely to report discussing medications (77% vs. 45%; p=&lt;.001), were more likely to answer knowledge questions correctly (risk reduction with statins 61% vs. 33%; p=0.07; knowledge about options 57% vs. 33%; p=0.002) and were more engaged by their clinicians in decision making (OPTION Score: 50% vs. 28%, difference 22% (95% CI 6.4, 36%), p=.01).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conclusion. Delete feasibility statement. Address what you found. Clinical outcomes were not different; knowledge and discussion in the encounter were diff between intervention and control. But tell the reader what this means. (see more questions below that might lead to some more explanation or reflection.) Acknowledge that the study was underpowered to show differences, but say what the direction of the two categories of results was and what you think it means.</td>
<td>We thank the reviewer for these suggestions. We have removed statements about feasibility from the Conclusion. The conclusion reflects the results in the following way:Page 2 Para 4 : “DAs improved decision outcomes without significant effect on clinical outcomes. DAs designed for point-of-care use with type 2 diabetes patients promoted shared decision making in nonacademic and rural primary care practices.”</td>
</tr>
</tbody>
</table>
| Background: Throughout the background and “comparison with prior research”, the authors refer almost exclusively to their own work. Are they the whole field? For this paper to relate to the general field requires identifying gaps in knowledge that this study is prepared to fill. Indicate where these results support prior research and where they don’t. | To our knowledge, there are no other trials of decision aids for diabetes for use during the encounter focused on the decisions tested in this trial. For context and comparison, we could have drawn a broader picture, but we believe shared decision making operates differently with different interventions and context making such indirect comparisons unhelpful. Our intent is not to be arrogant and exclusive, but rather to be pertinent and parsimonious. Indeed, our prior work took place in academic practices, and we wondered whether the results could reproduce in rural clinics. Also, we have compared our results with our prior research in the discussion section PARA 2 PAGE 15 as follows: “This study, along with others we have conducted, have found improvements in knowledge, decisional comfort, and patient participation in decision making, and little impact on adherence and other patient health outcomes. The Diabetes Medication Choice study [6] found a significant increase in knowledge over usual care (adjusted mean difference 1.10 (95% CI 0.11, 2.09), an increase of patient involvement on average of 21.8 of 100 points higher, but no impact on medication adherence at 6 months nor an impact on HbA1c levels. The first Statin Choice randomized trial [5] conducted in a specialty setting found that patients who received the decision aid were 22.4 times more likely to know their estimated cardiovascular risk than those in the usual care group, had greater decisional comfort (10.6 points higher on a 100 point scale), and better self-reported adherence at 3
months, odds ratio 3.4 (95% CI 1.5, 7.5). The second Statin Choice clinical trial [7] conducted by another group found an increase in patient knowledge of their cardiovascular risk as well (odds ratio 1.9, 95% CI 1.0, 3.8), an increase in decisional comfort among the informed subscale and support subscale, and no difference in medication adherence at six months. Our findings, while imprecise, are consistent with these prior results suggesting it is feasible to observe similar outcomes in academic and nonacademic practices, provided that the decision aids are implemented."

Page 3. Important claims are made that SDM is now required for clinical practice. Please provide a reference for the statement that CMS evaluates ACOs on ability to implement SDM.

We appreciate the opportunity to provide more detail in this regard. The Center for Medicare and Medicaid Services (CMS) evaluates accountable care organizations for their ability to implement SDM (Ref 10).

To make this clearer, we now have in PAGE 4 PARA 2 the following “While these trials provided evidence of efficacy of decision aids in patients with diabetes receiving care in academic clinics, little is known about the effectiveness of these tools in routine clinical practice, particularly in rural clinics. In this context of imperfect knowledge, state legislation in Washington and Minnesota and provisions in the Patient Protection and Affordable Care Act promote or require SDM [8,9]. The Center for Medicare and Medicaid Services evaluates accountable care organizations for their ability to implement SDM [10]. The National Quality Forum recommends measuring SDM as part of its framework for assessing quality of care in patients with multiple chronic conditions [11].”

This can be reviewed in the following website: Ref: Accountable Care Organization 2013 Program Analysis http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/sharedsavingsprogram/Downloads/ACO-NarrativeMeasures-Specs.pdf

Please provide a reference for the statement on pages 3-4 that the NQF recommends SDM in their quality of care framework.

There are two pertinent references, which we now include to support our statement in Page 4 Para 2:


As before…delete the feasibility statement We have removed the offending statement.
<table>
<thead>
<tr>
<th>On page 4 you say this is an effectiveness study. What was the measure of effectiveness, and what were the hypotheses?</th>
</tr>
</thead>
<tbody>
<tr>
<td>As above, we have now spelled out the goals of this trial in PAGE 5 PARA 2 as follows: “Thus, the objective of this study was to evaluate, in a cluster-randomized practical trial enrolling nonacademic and rural primary care practices and their patients with type 2 diabetes, the impact of patient decision aids vs. usual care on decision making measures, metabolic control and medication adherence.”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methods:</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is no section on trial design. Please add one</td>
</tr>
<tr>
<td>The trial design is addressed in the text using the following statements in the ‘Methods’ Starting on Page 5 Para 3</td>
</tr>
</tbody>
</table>

“We conducted a multicenter cluster randomized controlled trial set in nonacademic and rural primary care practices.”

**Participants**

Eligible participants were physicians, nurse practitioners, and physician assistants (i.e., clinicians) who cared for patients with type 2 diabetes at participating primary care practices. Practices were deemed eligible if they provided primary care for patients with type 2 diabetes. Minimal training was provided to clinicians that consented to participate [13]. Eligible patients were adults with >1 year of type 2 diabetes with a reason, identified by their clinician, to consider changing their antihyperglycemic or lipid-lowering regimens. For example, for the diabetes discussion, eligible patients had HbA$_{1C}$ > 7.3, were not using insulin, and were not taking >2 antihyperglycemic agents at maximum dose. For the discussion about statins, eligible patients should not be using statins and should not have contraindications for taking statins. All patients and clinicians signed written informed consent for participation.

**Approach and protection against bias**

Primary care practices were enrolled then matched by size ($\leq$ 2 clinicians or $>$ 2 clinicians) and randomly allocated by a statistician (who was the only team member aware of the composition of the
In it, please discuss the cluster randomized trial and the known problems and how you addressed them.

<table>
<thead>
<tr>
<th>In it, please discuss the cluster randomized trial and the known problems and how you addressed them.</th>
<th>As discussed above, we have now added a section to the limitations of this trial regarding the use of the cluster trial design in terms of number of clinics (clusters), difficulties with balance of participants, and potential problems with allocation concealment within clinics and selection bias in the discussion section.</th>
</tr>
</thead>
<tbody>
<tr>
<td>In particular, it is good that you provided the reviewers with the CONSORT checklist. However, more to the point is to use it to guide the write up. In particular, the diagram suggests this was a patient-level randomized trial, which it was not. Please show the centers and how many patients and clinicians were in each.</td>
<td>We have adhered to the recommendations of the CONSORT statement for cluster trials (<a href="http://www.consort-statement.org/extensions/designs/">http://www.consort-statement.org/extensions/designs/</a>) and have modified the figure to highlight that this was a clustered randomized trial.</td>
</tr>
<tr>
<td>Scattered throughout the report are various references to contamination. It appears that patients were either in the intervention or control group. However, clinicians were not. Please discuss this. How do you know contamination did not occur?</td>
<td>Contamination as discussed in the manuscript is assessed using the fidelity scale, which was created by the study team. This sought to ensure that clinicians would not recreate the decision aid when this was not available. However, this was nearly impossible as clinicians only had access to one decision aid (either the diabetes or statin choice) depending on their work site (there were no clinicians working in different clinics) allocation. If they were seeing a patient to discuss diabetes medications and were allocated to use the diabetes medication decision aid then they would have access to the tool to use it; if they were not allocated to use that tool then they did not have it available at the site and were instructed to treat patients according to usual care. This is an advantage of the clustered design. When we did look, we found that 20% of the content of the decision aid was mentioned in control encounters.</td>
</tr>
<tr>
<td>Among clinicians who had both intervention and control patients, what was their average OPTION score by condition?</td>
<td>There was only 6 clinicians who had patients in both control and decision aid that also had the encounters video recorded. For the statin discussion (11 (DA 8) encounters) the mean score was 39.9 SD (24.3). For the diabetes discussion (7 (DA=4) encounters) the mean score was 43.1 SD (29.4). The numbers are not sufficient to make any inferences on nor able to compare the values by arm.</td>
</tr>
<tr>
<td>You appear to believe the training was too brief. Do you know that from the OPTION score or the checklist? How does the checklist relate to the OPTION score?</td>
<td>We know this from the brief interaction with clinicians to train them on using the decision aid. The fidelity scores from the checklist suggest less than optimal use. Yet, this was sufficient to render a significant result. Thus, our impression is that some efficacy was not realized because of limited fidelity.</td>
</tr>
<tr>
<td>Please state in Table 1b how many physicians, how many nurses, how many PAs in each center delivered the DA.</td>
<td>We now provide this information in the Table reporting how many physicians and midlevel providers participated in each arm.</td>
</tr>
<tr>
<td>Was each of these video recorded and included in the OPTION analysis?</td>
<td>It is not clear to us what “these” refers to, but assuming it refers to the videos obtained, yes, all videos obtained were included in the OPTION analysis.</td>
</tr>
</tbody>
</table>

| Page 4 under participants. Please report the eligibility criteria more precisely. Patients with HgbA1c between X and Y? Cholesterol above? Not at guidelines of the clinic? What was the reason patients “had a reason to consider changing their regimens”? Were the meds they were on not consistent with guidelines? | Highlighted text is added to existing section:
Page 6 Para 1
Participants
“Eligible participants were physicians, nurse practitioners, and physician assistants (i.e., clinicians) who cared for patients with type 2 diabetes at participating primary care practices. Minimal training was provided to clinicians that consented to participate [13]. Eligible patients were adults with >1 year of type 2 diabetes with a reason, identified by their clinician, to consider changing their antihyperglycemic or lipid-lowering regimens. For example, for the diabetes discussion, eligible patients had HbA1C > 7.3, were not using insulin, and were not taking >2 antihyperglycemic agents at maximum dose. For the discussion about statins, eligible patients should not be using statins and should not have contraindications for taking statins. All patients and clinicians signed written informed consent for participation.” |

As this was framed as a practical trial, we did not want to exclude patients in whom clinicians thought that a discussion about medication change could take place. The constraints above were set to ensure there was still room for a discussion about medication choices, something we had to predict to enroll patients before the
consultations. Regarding cholesterol, we did not set a cholesterol threshold, as our decision aid is based on the randomized trials of statins in which a cholesterol threshold or target are not related to the risk reducing effects of these drugs, as documented in the Heart Protection Study, and discussions about statins are guided by the estimation of cardiovascular risk at 10 years. In summary, the trialists did not impose on the primary care clinicians a practice pattern or guideline that would have required or triggered the use of the decision aid.

<table>
<thead>
<tr>
<th>Data collection and outcomes: Whether or not there is more discussion of this information elsewhere, this manuscript has to be clear and easy to follow. Are there other measures you report (or don’t report) that were used? In particular, you have two main outcome categories: Quality of Decision Making and Clinical Outcomes. Please say which measures were used for each. Please group them that way in separate tables in the outcomes description and in the results.</th>
<th>We have tried to improve the consistency and clarity with which we report the outcome measures (for example in para 2, page 7). We must lean to some extent on our published work on this trial to report with enough detail. In some cases, we have opted to favor some degree of parsimony taking advantage of the publication of the methods paper in the open access literature, thus avoiding duplication. Yet, to increase clarity and detail we now report the outcomes in consistent categories as the reviewer notes. For instance, in the text under Data collection and Outcomes in the Methods section we had separated the outcomes into ‘Decisional Outcomes’ and Clinical Outcomes’ of which the results section reflects this and is in the same order. To maintain consistency, Table 5 has been re-labeled as ‘Clinical Outcomes – Lab results and medication adherence’. Table 2-4 have been adjusted so the label for the tables begin with ‘Decisional Outcome - …’ to address this.</th>
</tr>
</thead>
</table>
| • Effect (of what? By what measure? Table 3)  
  • Perception of Knowledge (by what measure? Table 3)  
  • Adequacy of Support (by what measure? Table 3)  
  • Knowledge transfer (by a knowledge questionnaire? How did you measure transfer?)  
  • Knowledge of Risk (what was the measure?)  
  • Decisional comfort  
  Self-report of a pertinent medication (By what measure? If yours, please include in the appendix)  
  • Level of patient engagement (There are two different Table 3s with two different measures. Which of the following is it? How did you use them? See specific questions below.  
  o Fidelity checklist? (Please include in Table 3 has been modified to change the labels so they now read:  
  Comfort of Decision (DCS)  
  Perception of knowledge (DCS subscale)  
  Adequacy of support (DCS subscale)  
  Effect (DCS subscale)  
  Knowledge Transfer was described in the data collection section as being 6 true/false questions about the patients options and the pros and cons of the agents. Knowledge of risk was also described in the data collection section as knowledge of the patients estimated risk of having a heart attack within the next 10 years with and without use of a statin.  
  We made the mistake of submitting two tables labeled as Table 3, but this has been corrected in this resubmission.  
  Clinicians’ effort to engage patients in decision making is reflected in the OPTION score and this is now explicitly |
| Table 3 has been modified to change the labels so they now read:  
  Comfort of Decision (DCS)  
  Perception of knowledge (DCS subscale)  
  Adequacy of support (DCS subscale)  
  Effect (DCS subscale)  
  Knowledge Transfer was described in the data collection section as being 6 true/false questions about the patients options and the pros and cons of the agents. Knowledge of risk was also described in the data collection section as knowledge of the patients estimated risk of having a heart attack within the next 10 years with and without use of a statin.  
  We made the mistake of submitting two tables labeled as Table 3, but this has been corrected in this resubmission.  
  Clinicians’ effort to engage patients in decision making is reflected in the OPTION score and this is now explicitly |
appendix.) How did this control for contamination? Did you adjust the results for those who met your minimum set? Were they all equally proficient? How is this checklist used analytically? Fidelity checklist? (Please include in appendix.)

- OPTION score?

linked.

Fidelity is described in the text as a checklist designed by the study team. The items have been added in the appendix. We evaluated all of the videos – fidelity scores turned into contamination scores for encounters in which the decision aid was not used. We did not adjust our results for those that met a minimum standard. Our sample size does not support further parsing of the data.

The data from the checklist is informative only, to let us know how clinicians are using the decision aid, what specific items they are addressing, and whether training was sufficient. We report findings in the manuscript as per our protocol..

• Decisional comfort (How did you use the Decisional Conflict scale to get this? Did you add up the three subscales? Are there norms for this group of subscales? If so, is this a clinically meaningful difference? You cite a validation paper. However, it is not clear that this relates to your 3 subscales. Do you have the results on the whole scale, but have reported selectively?)

An overall score for the three subscales is reported. As this score is not being compared to any minimum cut-off for decisional conflict, choosing to instead compare between arms and show the values of the individual subscales the study team feels the data is well represented and reflected accurately. We only collected the three scales and reported all information obtained.

OPTION score? You appear to use this throughout the report as a measure of patient engagement. It is a measure of physician behavior. Please correct this throughout. Or do you have another measure of patient engagement that I missed?

We have used a shorthand that we now see is inadequate. We now refer to the OPTION Score as a measure of: ‘clinician effort to engage patients in decision making’

In the report on clinician behavior, beginning on 10, you appear to report only the checklist. How did the intervention and control clinicians score on OPTION? As mentioned above, among clinicians with patients in both intervention and control groups, how do their scores compare on intervention and control encounters? There appear to be 19 clinicians who have patients in both groups, so this is a large number in the context of this study.

We now report both results in Table 3.

On page 7 you mention that the medication adherence was only considered among those who had a discussion of meds. Was this a post-hoc analysis after generating the OPTION scores? Is the discussion data only among a subset of patients in each group? This needs to be clear.

We analyzed all medication data that we obtained as reflected in the number of patients reported in table 5. To qualify for the medication analysis, we noted whether patients reported having had a discussion about the relevant medicines (rather than ascertaining this directly on the videos as suggested by the reviewer as this would have excluded too many encounters). This was
not a post-hoc decision, but rather a planned analysis per our protocol, and this was not a subset of all participants, but on the pertinent participants who had a medicines discussion in the consultation.

Please tell us what analysis measured the effectiveness of the study, your main focus. Specifically, how and why were the regression measures done and for what purpose? I don’t think I see them in the results tables.

Effectiveness can be judged from reviewing our results on knowledge transfer, reducing decisional conflict, increasing patient satisfaction, and impact on clinical outcomes; all of these results are fully reported. The language about the outcomes has been modified to clarify this issue.

<table>
<thead>
<tr>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please include denominators on each line of results tables. Denominators change with every line and are very hard to follow.</td>
</tr>
<tr>
<td>During the last review the study team addressed this issue. We added N's to every table and some in several spots or addressed in footnotes. We understand that the numbers change; unfortunately this is reflective of survey return rates and completion of all items within surveys, as well as the issues with patients returning for lab values or pharmacies returning records when requested.</td>
</tr>
<tr>
<td>The study team felt it was important to show all the data obtained rather than eliminating patients due to incomplete data. We felt this was the more complete approach to reporting our findings, even if it led to variable denominators per item.</td>
</tr>
</tbody>
</table>

Quality of decision making: patient outcomes. Were your hypotheses supported? If not enough power, please state the direction of the absolute results as well as the relative results for each variable that measures this.

The results section refers to the quantitative results for each of the outcomes; where necessary we have indicated whether the direction of effect is in favor of our hypothesis recognizing that some of the measures may not be immediately obvious to the reader.

<table>
<thead>
<tr>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>As noted, what you call patient engagement is based on clinician behavior. If you don’t have a validated measure of patient engagement, then a better term is needed for the clinician behavior. SDM in the encounter? Something accurate and clear.</td>
</tr>
<tr>
<td>As stated above, we have used a shorthand that we now see is inadequate. We now refer to the OPTION Score as a measure of ‘clinician effort to engage patients in decision making’</td>
</tr>
<tr>
<td>Can you report a level of SDM in the encounter from your OPTION data? Is that what you mean to report? So then does SDM in the encounter predict Quality of Decision Making (knowledge, medication discussion) or Clinical outcomes (adherence, HgbA1c, LDL cholesterol, starting meds, adhering to meds)? While underpowered, please report the absolute and relative results.</td>
</tr>
<tr>
<td>It is not best practice to look at subgroup effects or interactions between outcomes of a randomized trial. The proposal to relate OPTION scores with other outcomes falls into this approach which is not recommended. We now report the absolute results per arm, and comparisons in relative terms with appropriate measures of precision for each one.</td>
</tr>
<tr>
<td>Please do not report feasibility results on Statements about feasibility have been removed from</td>
</tr>
<tr>
<td>Page 12</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Please include a section on comparison of results with those of other DA trials in the literature, across conditions. What are the big take-home messages that this trial contributes?</td>
</tr>
<tr>
<td>Page 14 is the implications for policy. Can you say more about what you mean? Is it that:</td>
</tr>
<tr>
<td>o High-fidelity delivery of tools is challenging. You need better training and monitoring of implementation??</td>
</tr>
<tr>
<td>o Impact on adherence, HgbA1c and LDL is disappointing. But is this a function of the fact that few patients/clinicians made a med change? What was expected?</td>
</tr>
<tr>
<td>Question</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>What was the clinical problem that was behind the selection of patients? Out of compliance with guidelines? If so, is the challenge to get clinicians to recommend something different? To discuss with patients? If the decisions reached were shared ones, does this mean patients made an informed decision not to change meds? Or did the clinicians fail to bring it up? We need to know what you learned.</td>
</tr>
<tr>
<td>What do you mean that DAs promote patient-centered practice and patient engagement? Which measures? What insights have you gained? This is the heart of the matter, and it is not possible for an outsider to read between the lines.</td>
</tr>
<tr>
<td>A trialist should review the study to say if the authors have adequately reported this effectiveness cluster randomized trial.</td>
</tr>
</tbody>
</table>