Reviewer's report

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

Version: 3 Date: 8 May 2013

Reviewer: Margaret Holmes-Rovner

Reviewer's report:

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.

From the “lessons learned” 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. It basically shows no impact in clinical results because DAs did not impact clinical practice. Clinician and practice recruitment was very low. Selection bias is highly likely, but not acknowledged. 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. 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.
Mandatory revisions:

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. 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.

Abstract: Methods needs a trial design statement. 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) 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.

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.

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

As before…delete the feasibility statement to maintain the focus on effectiveness (p 4). On page 4 you say this is an effectiveness study. What was the measure of effectiveness, and what were the hypotheses?

Methods
There is no section on trial design. Please add one. In it, please discuss the cluster randomized trial and the known problems and how you addressed them. 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. 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? Among clinicians who had both intervention and control patients, what was their average OPTION score by condition? 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? Please state in Table 1b how many physicians, how many nurses, how many PAs in each center delivered the DA. Was each of these video recorded and included in the OPTION analysis?

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?

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. Decisional outcomes appear to be as follows. Is this correct? If not, please clarify in the manuscript. 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. Below is what I can glean from the article about outcomes and their measures:

- 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 (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?)
- 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.
  - Fidelity checklist? (Please include in 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?
  - 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
throughout. Or do you have another measure of patient engagement that I missed?

- 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.

Clinical outcomes:

- 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.

Analyses

- 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.

Results

- Please include denominators on each line of results tables. Denominators change with every line and are very hard to follow.

  - 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.

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

Discussion

- 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.

  - 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.

  - Please do not report feasibility results on page 12. Focus of this manuscript is effectiveness.

  - 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?
• Page 14 is the implications for policy. Can you say more about what you mean? Is it that:

  o High-fidelity delivery of tools is challenging. You need better training and monitoring of implementation??
  
  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? 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.
  
  o 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.
  
  o You say policy makers will not be happy with the clinical results. Is this something that can be remediated with better implementation? Better training? The abstract of your “lessons learned” paper is all about recruitment. In that case, should the policy makers discount these results until the recruitment issues are resolved? That is very important, if true.

A trialist should review the study to say if the authors have adequately reported this effectiveness cluster randomized trial.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare I have no competing interests.