Author's response to reviews

Title: Feedback GAP. Cluster-Randomized Trial of Goal-setting and Action-Plans to increase the effectiveness of audit and feedback interventions in primary care. Study protocol.

Authors:

Noah M Ivers (noahivers@gmail.com)
Karen Tu (karen.tu@ices.on.ca)
Jill Francis (j.fran시스@abdn.ac.uk)
Jan Barnsley (jan.barnsley@utoronto.ca)
Baiju Shah (baiju.shah@ices.on.ca)
Ross Upshur (ross.upshur@sunnybrook.ca)
Alex Kiss (alex.kiss@ices.on.ca)
Jeremy M Grimshaw (jgrimshaw@ohri.ca)
Merrick Zwarenstein (merrick.zwarenstein@ices.on.ca)

Version: 3 Date: 28 October 2010

Author's response to reviews: see over
28 October 2010

Dear Implementation Science Editorial Team,

**RE: Manuscript #:1004020503440057**
*Feedback GAP. Cluster-Randomized Trial of Goal-setting and Action-Plans to increase the effectiveness of audit and feedback interventions in primary care. Study protocol.*

The team is grateful for the positive feedback from Dr. Hysong and we thank Dr. Goderis for his thoughtful and thorough review.

Below, we respond specifically to each of the questions and concerns raised. To facilitate the editors in assessing our responses, we have taken the liberty of reordering some of the questions.

We have also attached a document with “track changes” and a properly formatted document.

We believe that our responses have satisfactorily addressed the concerns raised by Dr. Goderis and that the subsequent changes made to the manuscript have made it stronger. We thank you for the opportunity to improve our manuscript through this process.

Sincerely,

Noah Ivers
on behalf of the Feedback GAP study team.
Is the worksheet on paper or electronic and is it ‘ready to use’?
Is there an instruction manual to complete the worksheet?

The worksheet is on paper; as stated in the manuscript, it is delivered by post with the feedback reports. The prototype uploaded to the internet is smaller than the actual size of the worksheet making the write-in spaces appear smaller than in the intervention. We actively chose to try to limit the amount of materials sent to the participants and therefore did not send separate instruction manuals. On the worksheet, the instructions are to choose one target and set a goal for improvement in that target.

We did pilot test the intervention to ensure that potential participants would understand the instructions. We will add the following in the intervention section:

“We tested the worksheet design and all other intervention materials with a group of non-participating family physicians and they found it easy to use. Specifically, they reported that they found the instructions clear and advised no changes to the design.”

As reported in the protocol, the trial began in August 2010. In the trial, there is an instruction on the worksheet to fax back the completed version to the research team. In a brief review of the returned worksheets to date, it would appear that the instructions are followed without difficulty.

Some questions seem a bit trivial and large (e.g. if a patient comes to clinic and is not meeting targets, I will…). Not meeting HbA1c target will mostly induce treatment intensification and in some (or most) cases discussion about lifestyle; not meeting the BMI-target is totally different. And there are about 13 (or more) treatment targets to be met…. Other questions are not clear because they are extremely large: “for diabetes, I will improve”…: that can mean organizational changes (what possibilities), attitude changes, CME (on insulin therapy, new drugs like GLP-1 analogues,…). Is there any evidence that such large questions can guide FPs?

The intervention is supported by a broad range of theory and some empirical work in non-physicians, as described in the manuscript. We cannot be sure that this is the ‘best’ way of constructing a worksheet, but we did our best by involving an experienced multidisciplinary team and by piloting it with family physicians. The worksheet was specifically created to have participants complete action plans using “if, then” statements based on the data supporting the use of implementation intentions in achieving goals (see reference 15). The instructions on the worksheet specifically ask the participant to set the action plan in relation to the particular goal they have set for improvement. We recognize that this risks some participants either not completing the worksheet or completing it with a non-specific action plan, but this was a risk we felt worth taking. Furthermore, as described in the analysis section, we are planning an efficacy analysis by considering the extent to which worksheets were properly completed. Finally, we will be able to assess the use of the worksheet during the qualitative aspect. The following was added in the intervention section to clarify:

“Based on our review of the literature, the largest effects from goal-setting and action-planning seem to come from actually developing the plan (and linking it to a specific context to carry it out). For this reason, we chose not to provide participants with a list of possible actions. The participants, not the investigators, decide how to improve upon a care-gap which they identify as important.”
**Will there be some support to ensure (or to try to ensure) that FPs will fulfil their promises?**

There will not be further support. This is an insightful query and it is one of our underlying questions for the trial as a whole. Our literature review has shown that both goal-setting and action-planning are powerful interventions with respect to producing actual behaviour change, but the jury is still out with respect to their impact on health care providers.

We are also curious in future in understanding whether it is possible to get some of the positive benefits of academic detailing (aka educational outreach) without the cost. A few trials have compared academic detailing to feedback alone, and the effects seem to be positive, but inconsistent. Almost all academic detailing programs use performance feedback as a foundation for the intervention. It is possible that one of the major reasons that academic detailing is effective is that providers are assisted in setting particular plans in response to data. As mentioned in the discussion section, we believe that if our worksheet intervention is effective, it should then be tested against more intensive interventions like academic detailing where-in follow-up support is provided. If it doesn’t work, it may be necessary to test again in combination with other interventions.

In summary, we propose that the answer to this question should be found through a series of rigorous trials. We certainly hope that this trial will lead to future trials.

**Would it be useful to organize a group approach?** I can imagine that each single FP will not act on his own, but that they will ask the secretary to carry out the electronic search? But if a group approach will be organized, **what is the use of an individual (family physician orientated) worksheet?**

It might be very interesting to organize a group approach, but this would indeed require organizing as it is not the norm in Ontario for physicians to meet regarding quality improvement, even if they share an office-space. Certainly group approaches have been used before (e.g. ‘learning collaboratives’) but trials to date have not consistently shown that these approaches lead to improved outcomes. In future, we would be interested in formally testing whether adding group approaches to the basic design described in the current protocol is ‘worth’ the extra effort for both organizers and participants.

It seems like that the source of this question comes from a need for further information on our setting. Therefore, the following has been added to the study design section:

“In Ontario, family physicians do work in groups, but this generally involves sharing administrative resources not patients. The usual approach is for chronic conditions to be dealt with by the personal physician, while acute issues may be dealt with by whichever physician is available. Therefore, we believe that an intervention aimed at the physician rather than the entire group is appropriate.”

If it the purpose is to test the effect of the complementary tool (the worksheet) on Family Physicians, like it is stipulated in the study hypothesis, then the control group should compulsatorily receive the ‘usual’ feedback. This is part of the design and a consequence of the study question. **So there should not be a control group receiving no feedback at all. So why this excuse** in the sentence “The lack of a control group receiving no feedback at all
is both necessary (because participants expected something in return for contributing data) and pragmatic (because most quality improvement interventions include some degree of feedback, making this the ‘usual care’ comparator)?

We agree with this assessment. However, we also believe that many readers of quality improvement trials will expect to see an arm with usual care (no intervention). Head-to-head trials assessing different ways of conducting feedback were called for in the last Cochrane review of Audit and Feedback, but remain rare in the literature. It is for this reason that we have offered additional reasons to explain why we do not have an arm that receives no audit and feedback.

If the editors feel that we should change the manuscript, we would be pleased to do so.

The authors should specify that their method used is a mixed method, combining qualitative and quantitative research. (p. 7, study design).

The following change has been made:
“This is a mixed methods study built around a pragmatic, cluster-trial with two arms; one group will receive ‘simple’ feedback, while the other will receive ‘enhanced’ feedback.”

Explain the minimization procedure more in detail (e.g. what program used?). I do not understand that a minimization procedure can be non sequential since minimization is used to (prospectively) reduce imbalance.

As mentioned in the protocol, the software “MINIM” was used for the minimization procedure. The reference to this software is also provided in the manuscript.

Although we say that allocation is not sequential in the protocol, what we meant was that while allocation of clusters is sequential, the recruitment of clusters is not. In this trial, recruitment of the first fourteen practices (clusters) was completed prior to allocation. Then the allocation using minimization was conducted using study IDs instead of using names of the practices. We believe that this process substantially reduced the risk of bias since there was no chance that recruitment could be affected by allocation and it was difficult or impossible to know whether a practice was allocated to the intervention or usual care arm.

We have tried to make this clearer by changing the sentence “In fact, risk of selection bias is especially low in this study since allocation is not sequential.” to:
“We believe that risk of selection bias is low in this case because the recruitment of the first fourteen practices (clusters) was completed prior to allocation of the practices using the minimization software.”

P. 4: “A large gap between ideal and actual provided care”: there should be nuancing to this sentence. The gap between theory and practice is not entirely due to the providers’ performance. Patients’ characteristics and system characteristics are also associated with the quality gap.

We agree! The following was added:
“Thus, audit and feedback focuses on addressing the gap between ideal and actual care that is within the control of the health care provider and is often the foundation of multifaceted quality improvement interventions.”
I wonder how EMRALD can define diabetes when the EMR only uses free text.... What about the accuracy of laboratory and clinical data extraction from the EMR? Please give some details about this procedure prone to bias. There should also be a paragraph in the discussion section because it is a major weakness of the study, not only for generalizability, but also for internal validity.

While many EMR databases have not been adequately validated, nearly all of the data elements in EMRALD have been. The paper describing this process and the results is currently being submitted for publication but has not yet been published. Nonetheless we have published manuscripts describing the validity of our methods for identifying patients with diabetes and ischemic heart disease and have referred to these papers in the manuscript. In the section describing data collection, we have added further detail regarding their accuracy; our approach to identifying patients with heart disease and/or diabetes performs similarly or better than administrative databases or billing records.

This limits the external validity of the study only to the extent that other EMR databases will not have the exact same mechanisms in place. However, the knowledge to be generalized from the study is not dependent on the accuracy of EMRALD. If the project is able to offer a greater understanding of how feedback works and how it could be designed to work better, this knowledge will be useful regardless of the source of the feedback.

Finally, we respectfully disagree that any problems with data abstraction would produce a systematic bias and/or would limit internal validity of the study because there is an equal likelihood that data errors will occur in both the intervention and control arms. Furthermore, there is no reason to believe that the primary outcomes (including laboratory or blood pressure values extracted directly from specific tables in the EMR) would be affected by systematic bias.

Nevertheless, we agree that further explanation is warranted in the discussion section and the following has been added:

“Perceived data accuracy has been identified as an essential feature for acceptance of feedback reports. [van der Veer SN, de Keizer NF, Ravelli AC, Tenkink S, Jager KJ. Improving quality of care. A systematic review on how medical registries provide information feedback to health care providers. International Journal of Medical Informatics 2010 May;79(5):305-323.] and previous studies have shown that data quality from EMRs are uncertain [About quality of EMR-data: Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. BMJ 2003; 326(7398):1070]. While we have validated algorithms for identifying patients who have diabetes and/or IHD in this database, the automated chart-audits that are the basis for both the intervention and the outcome assessment may occasionally have errors. However, there is no reason to believe that any such errors occurring in the data abstraction process represent a risk for bias; any problems with the algorithm would be equally likely to occur in either arm of the trial.”

P. 8: Please describe the setting compared to the general setting in Ontario.
Also in the table, please compare the participating FPs with the general FP population. This is important to have an idea about possible selection bias and to have an idea in what setting the trial was carried out. “However, this concern is partially mitigated by the varied characteristics of physicians in the sample.” Why is this so? Please explain / give some evidence;

We have added the following to clarify some relevant aspects of the setting:
“Setting

The Ontario Health Insurance Program pays for doctor visits and laboratory tests, but covers medications only for the elderly or those on social assistance. Over half of the primary care providers in Ontario have eschewed the old model of fee-for-service and joined primary care reform models where-in capitation plays a large role in compensated for patient care. To earn the capitation fees, physicians and patients must co-sign an agreement that officially adds the patient to the physician’s roster; through this process, patients are encouraged to seek care primarily with their own provider. Only data from ‘rostered’ patients are included in the trial.

All the physician participants in this project roster their patients and most also benefit from the newest primary care reform process which provides funds to hire allied health care providers to work in the clinic. Although less than half of Ontario family physicians use EMR, Practice Solutions® EMR has 45% of the Ontario EMR market. [OntarioMD funding eligible EMR offerings: Vendor market share [homepage on the Internet].: OntarioMD Inc. [updated July 30, 2010; cited September 9, 2010]. Available from: https://www.emradvisor.ca/compare]

Unfortunately, we do not have ready-access to further data that would allow us to compare the participating providers to the ‘average’ Ontario family physician, nor do we have data regarding the ‘average’ Ontario family physician that is using EMR.

As described in Table 1, the providers vary widely with respect to their experience and their location in the province. Half of the participating providers are female. In this way, we have shown that there is some variety in the characteristics of the participants. We agree that this does not remove concerns regarding selection bias – we have stated plainly that the participants were found through convenience sampling and are not necessarily representative – but we do believe that the variety in location, gender, and years experience partially reduce these concerns, as stated in the manuscript. Finally, the issue is again revisited in the discussion where is it described as an important limitation with respect to generalizability.

P.14: 13.7 (=14) clusters ; how many physicians will be involved?

In table 1, we have shown the characteristics of the initial 54 participating physicians who have 4593 patients with either heart disease or diabetes.

What about possible contamination by patient flow? Are patients assigned to one FP/clinic or is there a freedom of choice? If yes, what expected impact on the trial?

This question seems to arise due to our incomplete description of the setting. We have added an explanation explaining the process of building a practice ‘roster’ and that only ‘rostered’ patients will be included in the trial. Although they can switch if they choose, it is unlikely to occur. Although the intervention is directed at the provider, allocation was conducted at the level of the clinic to reduce the chance of contamination.

P.16 To what degree the governments feedback will or will not interfere with trial?

The following has been added in the qualitative section:
“While the ongoing government feedback will likely enrich the qualitative component of the study, we do not believe that it will impact the inferences made from the trial. (All participants will receive the government feedback, but the government feedback does not explicitly encourage goal-setting or action-plans.)”