

Author's response to reviews

Title: Rational Prescribing in Primary Care (RaPP): Process evaluation of an intervention to improve prescribing of antihypertensive and cholesterol-lowering drugs

Authors:

Atle Fretheim (atle.fretheim@nokc.no)

Kari Havelrud (kha@nokc.no)

Andrew D Oxman (oxman@online.no)

Version: 2 **Date:** 12 July 2006

Author's response to reviews: see over

Submission of revised manuscript:

1450216919103765 Rational Prescribing in Primary Care (RaPP): Process evaluation of an intervention to improve prescribing of antihypertensive and cholesterol-lowering drugs

Dear Editors.

Please consider the revised version of our paper for publication in Implementation Science.

We have responded to every comment from the two reviewers point by point below, and we you will be satisfied with our responses.

Oslo, July 12th, 2006

Regards,
Atle Fretheim (for the authors)

1st Reviewer's report

Title: Rational Prescribing in Primary Care (RaPP): Process evaluation of an intervention to improve prescribing of antihypertensive and cholesterol-lowering drugs

Version: 1 **Date:** 30 May 2006

Reviewer: Alicia O'Cathain

Reviewer's report:

General

I want to see process evaluations undertaken and reported and I would like to see this one published. It is a large quantitative process evaluation which allows one to address some important questions. However I think this paper needs further clarification.

This is a predominantly (or totally?) quantitative process evaluation focused on the intervention arm of an RCT. It is important to explicitly inform the reader of this to distinguish it from qualitative process evaluations, especially since the point you are making is that a qualitative approach would have been helpful.

Response: The last sentence of the Background in the abstract now reads “We carried out a predominantly quantitative process evaluation to help explain and interpret the trial-findings.”

The analysis undertaken here can explain variation in size of effect in the intervention arm of a trial. You had variation for the effect size in the prescribing outcome (Fig 1) but you say there was little variation for change in achievement in treatment goals (Fig 2). In this latter case you are attempting to explain very little variation and it is important to be clear about this.

Response: We state very clearly, in both text and figures, that variation in change was much greater for the prescribing of thiazides than for achievement of treatment goals, and since the analysis is about explaining variation, we believe this should be sufficiently clear.

Finally, you must make it clearer that you are not trying to explain change in assessment of cardiovascular risk because you only have post intervention data rather than effect size. You say this early on in the paper but are not clear that the variation you then try to explain is variation in assessment of cardiovascular risk. I think it is confusing to have this reported alongside effect sizes. It is too easy to think that you can say something about the trial from this data when you are really saying something about the state of assessment of cardiovascular risk in general practice.

Response: We agree that the outcome data for risk assessment are weak, due to the lack of baseline measurements. We wish to keep the data in the analysis, but have made several changes in response to this fair criticism:

a) We now point out the lack of baseline-data under Methods (last paragraph):

“For two outcomes (prescribing of thiazides and achievement of treatment-goals) we had measurements from before and after the intervention, and we used difference as the dependant variable. For assessment of cardiovascular risk we only had post-intervention data, which we used as the dependant variable.”

b) We also point this out under Results (Analysis-subsection, 1st paragraph):

“The degree of change in thiazide-prescribing varied considerably across practices (figure 1), while the change in achievement of treatment goals was more uniform, and in most cases close to zero (figure 2). There was wide variation across practices on the extent to which doctors were using cardiovascular risk assessment tools. Proportions ranged from zero to 100 % (median 5 %, mean 17 %). We did not have baseline measurements for this outcome, thus we could not estimate whether there had been a change in performance from before to after the intervention.”

c) and in the Discussion (5th paragraph):

“Another weakness is that we did not have baseline data for the outcome measure for use of risk assessment tools, which meant that we could not estimate change in performance in relation to the intervention. Whether the post-intervention rate is a valid effectiveness measure is highly questionable.”

d) We have also separated the multivariate regression analysis for this outcome from the two others (there are now two tables, Tables 5 and 6, instead of one)

Also, I didn't really understand why you only had post intervention data when your trial results say that there was no change in this outcome.

Response: In the RCT we compared post-test data in the intervention and control groups for this outcome. Thus, we agree that we should avoid using the term “change” when describing the effect on this outcome. We had only used this term in the Background section of the Abstract, which now reads:

“A randomised trial of a multifaceted intervention for improving adherence to clinical practice guidelines for the pharmacological management of hypertension and hypercholesterolemia increased prescribing of thiazides, but detected no impact on the use of cardiovascular risk assessment tools or achievement of treatment targets.”

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Abstract

It would be helpful to state that it is a predominantly/totally quantitative process evaluation. “Semi structured interview” implies qualitative research.

Response: Done (see above).

Was this a largely structured telephone interview with an open question?

Response: we have expanded the description of the telephone interview:

“Within three months after the outreach visit we conducted semi-structured telephone interviews with physicians in the intervention-practices, asking about how the intervention was perceived and their attitudes towards the recommendations we were trying to implement. The response options were “yes” or “no”, or on a 3-point scale (usually “negative”, “neutral”, “positive”), followed up by an open question, such as “Why?”.”

What is the status of your univariate and multivariate analyses? Why do you report both in the body of the paper yet just report the multivariate here? This is important and relates to the point I make later about the whole analysis.

Response: We have edited the Analysis sub-section under Methods, in order to better explain our analytical approach: The univariate analysis was planned as (and still is) an initial analysis in which we try to identify the most appropriate variables to include in the main, multivariate regression model (s) - one per outcome:

“We selected potential explanatory variables for each main outcome based on our own judgement and discussion with a general practitioner. We conducted initial univariate regression analyses to explore the association between the selected variables and the observed variation across practices. The variables that predicted the dependant variable at a statistical significance-level of $p < 0.30$, were included in the main analysis, which was a multivariate regression model for each main outcome.”

“we promoted” sounds like it was an intervention you developed. If it was you, you need to say this because the GPs may have been keen to offer you positive comments as developers of the system.

Response: Yes, we developed the intervention, and we now say this in the paper. Background (1st paragraph):

“We tested the effectiveness of a multifaceted intervention we had developed for improving adherence to clinical practice guidelines for the pharmacological treatment of hypertension and hypercholesterolaemia.”

And in the Discussion (2nd paragraph):

“our multifaceted intervention was specifically tailored to target barriers to change, including attitudes”

Also, we have added a comment on the risk there was that the GPs were eager to please us (Discussion, 3rd paragraph):

“may be partly due to social desirability bias: The interviews were conducted by a member of the research team; often by the same pharmacist that had visited the practice a few months earlier.”

The last sentence in the conclusion should be removed and left to the discussion.

Response: OK, done.

Background

Where did your hypotheses come from? Sometimes hypotheses emerge from a process evaluation which can then be tested in the outcomes data (see Oakley BMJ 2006).

Presumably you developed these hypotheses prior to the process evaluation?

Response: Yes, they were mainly developed at an early stage, when we decided what to include in questionnaires etc. We have inserted a “had” to make this more clear (Background, last paragraph):

“We decided prospectively to carry out a process evaluation of the implementation of the intervention.....We had hypothesised that the impact of the intervention would be correlated to several variables, including practice specific factors such as the attitude among the physicians towards the recommendations, and process-measures, such as the proportion of physicians attending the educational outreach visit.”

Analyses

You could make it clearer that you are looking for explanatory factors for variation in outcomes.

Response: We have added a sentence about this at the end of the Background-section:

“Thus, the main objective of this analysis was to identify factors that could explain variation in outcomes across practices.”

Results

I think you really need to consider the status of the univariate and multivariate analyses. The multivariate analysis will suffer from missing cases, and unless you do something about this (like data imputation) you may be undertaking these analyses on 50 cases rather than 70+. I would like you to consider how many cases you based these multivariate analyses on and clearly display this.

Response: We agree and have added information about this in the paper. The basis population of 66 practices is described (Results, Regression analysis sub-section, 2nd paragraph):

“The basis population for all analyses included the 66 practices for which we had completed questionnaires by pharmacists.”

And in Tables 5 and 6 we report the number of included practices (e.g. “N=61”). Thus, we have not done any data imputation.

Also you may want to explicitly consider the power of the multivariate analyses.

Response: Estimating the power to detect the various observed sizes at a statistically significant level is a complex undertaking in multivariate analyses. We could more easily estimate the power to detect a significant R^2 , but explaining the variance is not really what we are after here. Thus, we prefer not to enter into a power-estimating exercise. Also, we have commented on the lack of power in our study in the Discussion (8th paragraph):

“we may simply have had too little statistical power to detect important factors.”

By privileging the multivariate analyses you ignore in your abstract and results a number of useful findings from the univariate analysis. For example, the fact that a smaller effect was found if less GPs attended training - you make something of this finding in the discussion. I would really like you to think more about your analysis and the conclusions you can draw from it.

Response: We view the univariate analysis simply as a preliminary step in establishing the multivariate model, i.e. our main analysis. We do not want to put emphasis these initial findings, which we consider potentially misleading. For instance, the effect-sizes can change dramatically when adjusting for the other variables in the multivariate model.

Regarding the number of GPs attending the educational session: Our mention of this in the Discussion refers to the multivariate models. And we have made this explicit by adding “(tables 5 and 6)” in the text.

95% CI around key findings would be useful.

Response: Agree, done.

Discussion

When researchers undertake in depth interviews they are aware that interviewees give accounts, like “I am a good doctor”. There is no reason why they would not do the same in a survey. You have GPs claiming to do risk assessment but were found not to be doing it. What other claims did they make that you may wish to be wary of?

Response: We have used the outcome with the largest discrepancy between what doctors reported doing and what we found from medical records data (risk assessment). This was for illustrative purposes, and we don’t feel that we need to report similar comparisons for the two other outcomes (i.e. 35% of doc’s reported using thiazides as 1st choice, however thiazides were only prescribed in 17 % of cases; or that 71 % of doctors thought that most of their patients achieved treatment goals, while we found that only 32 % of the patients actually did). We could report this too, perhaps, but this is not the main focus of the paper.

Some points in the discussion related to limitations of the RCT not the process evaluation and these should be removed e.g. less than half practices invited agreed to participate, the last two paragraphs in discussion.

Response: Agree. Done.

Table 4

A table of p-values does not communicate very much (Table 4). Could you report effect sizes as well as p-values?

Response: In line with our accounts for why we privilege the multivariate analysis, we do not want to report the (potentially misleading) effect-estimates from the univariate models. We do agree that a table of p-values does not communicate very well, however the table serves the purpose of documenting how we arrived at our multivariate model, and it gives the reader an opportunity to see how close some of the excluded variables were to being included in the model.

What is the pharmacist dummy variable in Table 4? Is this the three pharmacists who trained GPs? You could make this clearer. Also, the intervention had a different effect depending on the pharmacist – how interesting.

Response: Yes, this is the (four) pharmacists that visited the practices. We have now included information about this in the Background (2nd paragraph):

“The intervention was multifaceted. It included an outreach visit conducted by one of four pharmacists recruited and trained specifically for this purpose.”

We do not want to put too much emphasis on the observed difference associated with which pharmacist was doing the visit because:

- a) none were found to be statistically significant (p-values, 0.89; 0.23; and 0.11)
- b) the effect-size is very small (at least for two of the three pharmacist variables included in the models)
- c) the finding is inconsistent across outcomes

When you included doctor’s attitude towards printing reminders etc in the analyses, no doctors were negative about this. How was it dealt with in the regression?

Response: We have used the numerical values for each response category (found at the top of Table 2). We are aware that we lack negative responses for many of these variables, and that the models that include those variables may not be generalisable to practices with more negative attitudes. We have therefore added this in the Discussion, 6th paragraph:

“The attitudes among doctors, as perceived by the pharmacists, were rarely negative, Thus, our models are not necessarily applicable to practices where more negative attitudes dominate.”

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Discretionary Revisions (which the author can choose to ignore)

Using the difference between two time periods is not as efficient as analysis of covariance where you use the last time period as the dependent variable and adjust for the first time period before testing independent variables. You may wish to check whether your findings are similar using this alternative analysis.

Response: We have reflected on this, and we have concluded that we wish to stick to our original analysis and not add another one. Analysis of covariance could be done, but there are also problems with that. For example, using post-level as dependent variable could underestimate effect-sizes (i.e. level of change) when baseline levels are low (due to ceiling effects, i.e. lower level of adherence at baseline can be expected to be associated with a higher degree of change, and vice versa). There are certainly ways of coming around this, e.g. by stratifying practices according to baseline level, but we believe that our approach of using change as the dependant variable and considering baseline level as a potential explanatory variable (it is included in one of our multivariate models) makes more sense – at least it does to us!

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

2nd Reviewer's report

Title: Rational Prescribing in Primary Care (RaPP): Process evaluation of an intervention to improve prescribing of antihypertensive and cholesterol-lowering drugs

Version: 1 Date: 15 June 2006

Reviewer: trudy van der Weijden

Reviewer's report:

Fretheim et al. Rational prescribing in primary care (RaPP): Process evaluation of an intervention to improve prescribing of antihypertensive and cholesterol-lowering drugs.

General

- Prospective process-evaluation of a large-scale implementation study in primary care (73 Norwegian general practices in the intervention arm!). The results of the implementation strategy were lower than expected.
- The intervention: outreach visit including feedback, electronic reminders integrated in the registration system, software to assess CVD risk and print patient education materials.
- This is a relevant exercise as we still lack knowledge on the explanation of unexpected low results. The high number of participants enables the authors to run explorative multivariate models.
- Well-written paper by a well-known group of experts. It is a peculiar finding that the one outcome that was found to show effect in the trial - prescription of thiazides - was in this process evaluation found to be the only factor that GPs were really negative about. Why did the authors not refer to their earlier work on this? (Fretheim BMC Health Services Research 2004).

Response: We have added it (Discussion, 2nd paragraph):

"However, doctors' attitudes had been identified as a likely reason why it could be difficult to increase the use of thiazides, and our multifaceted intervention was specifically tailored to target barriers to change, including attitudes [3]."

- The approach of this process evaluation was not particularly innovative. The message that more in-depth qualitative research should have been done is not new. There do not seem to be interesting learning points in the conclusion.

Major revisions

- The number of null hypotheses is limited. I miss for example: The impact of the intervention might have been correlated by the GPs' perception of the pharmacists' status as an opinion-leader. Is the pharmacist used as a variable?

Response: There were four pharmacists that conducted outreach visits, and we did test for association between which individual pharmacist conducted the visit and the three dependant variables (table 4, "Pharmacist dummy-variables")

- pag. 5: the pharmacist scoring on his/her impression of how the physicians of the practice reacted was apparently (looking at the numbers in table 2) expressed for the group (in case of duo or group practices) instead for the individuals. This is a potential weak point. The GPs' attitude might differ between colleagues (this is better dealt with in table 3) and the pharmacist might record the attitude of the most dominant GP only. Above that, it is not self-reported attitude but attitude perceived by the teacher who has tried to improve their attitude. So, measurements might be invalid.

Response: We have mentioned the problem of validity of measurements in the Discussion (8th paragraph):

"Many of the explanatory variables we have used are based on pharmacists' impressions during outreach visits or self-reporting from physicians and these may be inaccurate."

I find the ratings on "their attitude towards you" and "how do you rate your own performance" a bit strange.

Response: We were not surprised by these ratings.

- The 3 outcomes:

1) assessment of CVD risk. Might the negative result of the trial be caused by GPs' underreporting on this? He/she did not register if the risk was assessed before prescription.

Response: This was measured by asking doctors per telephone about specific patients that had been started on drugs during the study period. We believe that the doctor would have informed us during the interview if he/she usually estimated risk, but did not find a note of it in the medical record (in most cases we interviewed the doctors who had actually prescribed the medication). Also, to fulfil this outcome-criterion the doctor had to have done a formal quantitative estimation of risk (using a calculator, scoring-table etc). We believe doctors that did this would usually make a note of it in the medical record, and also: We do not find it very surprising that formal risk assessment was done in less than 20 % of cases.

We have added a description of how this information was collected (Methods, Data-collection, last paragraph):

"Data on prescribing and achievement of treatment goals were extracted from the electronic medical records. The data on prescribing enabled us to identify patients that had been started on medication, and we asked physicians per telephone about whether they had conducted cardiovascular risk assessment first."

2) the achievement of treatment goals. In the core paper on RAPP I have read that it was measured at patient level (patient outcomes on cholesterol and blood pressure) and not by performance of the GP (e.g. increasing drug dosing). So the outcome is also dependent on the adherence of the patients and failure of the strategy is not only explored by process evaluation of the doctors' adherence to the intervention protocol.

Response: Yes, we agree, and this is addressed in the Discussion (9th paragraph):

“achievement of treatment goals probably depends as much on patient behaviour as on the actions taken by doctors.”

- Confusion: In the introduction the number of 146 practices is mentioned. In the results I found the number in the intervention arm, n=73, that is subject to this process evaluation. But at page 9, second paragraph, the control practices suddenly pop up. What is going on?

Response: We agree that this was both confusing and unnecessary, so we no longer mention the control practices.

Minor essential revisions

- Why was the electronic reminder and the CVD-risk calculator not given more attention, e.g. by evaluating the frequencies of using the system? It is an essential part of the intervention.

Response: Our system did not allow for collecting such data.

- pag. 6: the fact that pharmacists who conducted the outreach visits performed the telephone interviews might be a weak point, in the responders being too socially desirable in their answers. Should that not have been a more independent interviewer?

Response: Possibly. On the other hand, getting through to doctors on the telephone is a challenge and we believe(d) that the chance of succeeding was better if the call was made by people to whom the practices had some form of relation. We have added this point, though, in the Discussion (3rd paragraph):

“This discrepancy may be partly due to social desirability bias: The interviews were conducted by a member of the research team; often by the same pharmacist that had visited the practice a few months earlier.”

- pag 6/7: univariate analyses were conducted and variables with significance level $<.30$ were included in the multivariate model. I do not understand the sentence in between on "selection based on our own judgment and discussion with a GP". Do you mean that next to the somewhat significant variables, other variables could also have been included even if they were far from significant in the univariate analyses? Confusing.

Response: We have edited this section in an attempt to clarify what we did (Methods, Analyses sub-section, 1st paragraph):

“We selected potential explanatory variables for each main outcome based on our own judgement and discussion with a general practitioner. We conducted initial univariate regression analyses to explore the association between the selected variables and the observed variation across practices. The variables that predicted the dependant variable at a statistical significance-level of $p < 0.30$, were included in the main analysis, which was a multivariate regression model for each main outcome.”

- pag. 8: feedback from physicians. An estimated 195 physicians. Is that correct? There were 70 practices with on average 2.3 GPs = circa 160 GPs I should think.

Response: The 2.3-figure is the average number of physicians attending the outreach meeting (85% attendance rate), while we wanted to interview all doctors in the practices (also non-attendees). Adjusting for this, the numbers fit quite well (85 % of 195 = 166).

- pag 11: the external validity of the study is not relevant in a process evaluation.

Response: Agreed – we have taken this out.

- pag. 11: the 3rd paragraph needs more text or otherwise might be left out. The authors refer to their publication on the OFF-theory, a both serious and hilarious statement on their disbelief in the value of theories on behaviour change. The info has a high degree of insiders information. It does not mean anything to the reader who is not aware of the Oxman-Eccles debate.

Response: We agree that one has to be an “insider” to understand this paragraph in full. However, it is short, keeps the debate going, and is – we think – amusing.

Discretionary revisions

- pag 5: the pharmacists completed a questionnaire: refer to table 2

Response: OK.

- pag. 6: the telephone interviews after 3 months, refer to table 3.

Response: OK.

- pag. 6: the sentence on the small financial compensation seems completely lost in between the other sentences on the telephone interview.

Response: We agree, and have moved it to the end of the paragraph.

- PAG. 8: bottom of page How could it how have been more useful?

Response: Corrected, thank you.

- fig. 1 and 2 seem not necessary.

Response: We would like to keep them – we think they illustrate the variation in change in a good way.

- pag. 12: last sentence: with = to have?

Response: Yes, we have changed it.

- pag. 13: "The authors observed which is consistent with our own observation." I must have missed something? Where did you report on this observation that the effect was smaller in practices where the pharmacist was unable to meet al doctors?

Response: This is found in Tables 5 and 6. For two outcomes (change in thiazide-prescribing and rate of risk assessment) we found such an association, although not statistically significant (p=0.13 and p=0.11).

- pag. 21: 7th item: of

Response: Corrected, thank you.

- pag 25: 3rd dependant: achievEment

Response: Corrected, thank you.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.