

## **Author's response to reviews**

**Title:** Implementing Electronic Clinical Reminders for Lipid Management in Patients with Ischemic Heart Disease in the Veterans Health Administration

### **Authors:**

Anne E Sales ([anne.sales@ualberta.ca](mailto:anne.sales@ualberta.ca))  
Christian D Helfrich ([christian.helfrich@va.gov](mailto:christian.helfrich@va.gov))  
Michael P Ho ([michael.ho@va.gov](mailto:michael.ho@va.gov))  
Ashley Hedeem ([ashley.hedeem@va.gov](mailto:ashley.hedeem@va.gov))  
Mary E Plomondon ([meg.plomondon@va.gov](mailto:meg.plomondon@va.gov))  
Yu-Fang Li ([yufang.li@va.gov](mailto:yufang.li@va.gov))  
Alison Lucarotti ([alison.lucarotti@ualberta.ca](mailto:alison.lucarotti@ualberta.ca))  
John S Rumsfeld ([john.rumsfeld@va.gov](mailto:john.rumsfeld@va.gov))

**Version:** 2 **Date:** 22 July 2007

**Author's response to reviews:** see over

Editor's comments (Brian Mittman):

I endorse all of the reviewers' comments although I am less concerned than Reviewer 2 with the adjusted results you report, in which you adjust for reminder use. I recommend additional text to explain why you performed this adjustment and why the study was not completely "clean" in terms of access to, and use of, reminders. Do you have data on the timing of reminder activation in the 3 intervention sites and the two comparison sites that activated the reminders? Because the "intervention" consists of the reminders plus the efforts to encourage and facilitate their use, more detailed information on exposure would help explain the value of the adjusted and unadjusted data.

*We appreciate the reviews, and your comments helping to evaluate required changes in the paper. We have attempted to respond appropriately to all the comments, as described below, and where changes were either not possible or appeared inappropriate given the details of our study, we have noted this in our response. We have used **bold** font for all changes to make them easier to identify in the revised text.*

*We recognize that we failed to be clear about the intent of the paper in the version originally submitted for review. We have attempted to clarify that the focus of this paper is not to evaluate the effectiveness of electronic clinical reminders, but instead to report on an implementation effort, consisting of facilitation efforts and identification of clinical champions at each intervention site to champion the use of clinical reminders after they had been released nationally, installed, and activated at each site.*

*We have also attempted to clarify that although we initiated the study with intentions of conducting a lagged quasi-experimental study, we were not able to complete the lagged design, and because of delays in the release of the reminders nationally and other issues, the resources available to complete the study were depleted before we were able to initiate the intervention in the comparison sites. We have also attempted to clarify why we did not randomize the sites. Although we were interested in attempting to study the effectiveness of the intervention, the reminders were very new, and we felt it was premature to insist on randomizing sites. Because of many issues that arose because of the delay in release of the reminders, we were not able to initiate much contact with the comparison sites, until the provider survey which was fielded towards the end of the intervention period. It was at this point that we learned that the comparison sites had a history of using reminders systematically over a long period of time, information which was somewhat at odds with the information we had received initially from regional leaders.*

*Finally, we note that we were breaking new ground within VHA for many of the measures included in our study. The capacity to develop national reminders was new, and together with one other group, the Mental Health Quality Enhancement Research Initiative, we were among the first to attempt to develop and use reminders for clinical quality improvement. Getting approval at the national level to access the programmers and other personnel required to develop the reminders, first, then test and deploy them, was a two year process, which we did not anticipate. It was particularly delayed because no procedures were in place as we initiated the project, although we received considerable encouragement from the national Research and Development offices. We assisted in developing the necessary procedures. After deployment finally happened, we found ourselves in the position of assisting sites to use the new software in ways we had also not fully anticipated, particularly in the generation of reports using the reminder software. We had anticipated that a national process of reminder reports, which would be sent to a national database, would accompany the initial clinical reminders. This latter*

*process, however, was not coupled with the clinical reminder development, and did not occur until over two years later, long after the project was completed. As a result, one of our key process measures, the reminder reports, was not available without considerable effort by the sites, which we had not anticipated or expected.*

*We have added details about the timing of the reminder activation in the text on page 8. Unfortunately, we did not have information about when the reminders were activated in the comparison hospitals, although our understanding from discussions with regional leaders suggests that it was not until at least 6 months after the intervention began. However, we do not have firm dates, so stated that we are unable to time the activation of the reminders in the comparison hospitals.*

Please review the manuscript carefully from the perspective of non-VA and non-US readers. Terms such as "patch" (final paragraph on page 6) and "reminder reports" require explanation or replacement. Additional minor suggestions include a suggestion to specify the number of satellite clinics per hospital (see "study design" in the "methods" section) and to explain the Seattle and Denver intervention teams (see page 6).

*We have made wording changes as suggested on page 6 and page 11 to change terms that are not clear. We have also included the number of satellite clinics in the study design section and explained the Seattle and Denver teams as suggested.*

Trudy van der Weijden comments

The effect of the intervention has been described for the actual application of the reminders. The effect was also analysed with controlling for the degree of reminder use as it was self-reported by the professionals. This is in my view a serious flaw. Providers generally overestimate their performance, and probably also their use of educational or quality assurance strategies to which they are exposed.

*We appreciate this comment, and agree with the reviewer. We have added text on page 14 to emphasize this point.*

But my main criticism is that we should not correct for actual reminder use, as the fact that reminders have not always been used is real practice and an important finding in itself. (We would e.g. not believe results from a drug trial that were blown-up through correction for patient adherence to the medication.)

*We have added text throughout the paper to make the point more clearly that we are not reporting on the effectiveness of reminders. Instead, our main purpose in this paper is to report on the impact of the intervention to attempt to improve use of reminders. In this context, we believe it is appropriate to adjust for actual reported reminder use. We agree that there are difficulties with this analysis, and we have attempted to be careful in our language to suggest impact rather than state emphatically that we believe that we present strong evidence.*

*However, we feel it is important to make it clear that we do believe that it is appropriate, given the intentions of our study and of this paper, to include a term for reported reminder use in estimating the impact of reminders on patient-level outcomes. If we were attempting to provide estimates analogous to those in clinical trial designed to estimate the effect of a drug therapy on patient outcomes, then it might be appropriate only to use an intent-to-treat analysis. However,*

*as noted in Shadish, Cook and Campbell (Experimental and Quasi-Experimental Designs for Generalized Causal Inference, Houghton-Mifflin 2002) on page 320, while an intent-to-treat analysis does yield unbiased estimates of the effect of randomization to treatment or not, it does not yield unbiased estimates of the effect of the actual intervention. We did not randomize, so an intent-to-treat analysis does not make sense in our context. We believe that by presenting results with and without adjustment for the reported use of reminders, we are giving our readers more information than if we made a decision to use an intent-to-treat analysis, or its analogue.*

- Why was the intervention not randomly allocated between the western and eastern side of the Rocky Mountains networks?

*We have added text on page 6 to address this and related points. We were not attempting to test the effectiveness of the reminders. The project was initially designed as the first effort in developing a complex intervention to support implementation of the reminders in a larger group of hospitals. However, due to the delays noted in the text and in the response above to Dr. Mittman's comments, we did not have an opportunity to continue development of the intervention and expansion to new sites within the VA. It is clear from this comment as well as those of the statistical reviewer that we were not sufficiently clear in our description of the project, and we have attempted to clarify more carefully to distinguish between this study and its design and studies that attempt to determine the effectiveness of an intervention.*

- No power calculation seems to have been done.

*Given that we were not attempting to estimate the effectiveness of reminders, we did not perform power calculations. We do provide intra-cluster correlation estimates at the bottom of Table 4 for the correlation among hospitals for the two patient-level outcomes that we report.*

- Data on the number of reminders due and reminders satisfied were available only for the intervention arm, for the latter half of the intervention period.

*This is correct. Reminder reports were only available if the reminders were installed and activated. In addition, creating the reports requires additional work locally at the sites, and could not be done by study personnel due to restrictions on access to the local clinical databases. As a result, we were dependent on site personnel to create the reminder reports, and time and availability of personnel with the expertise to do this were in short supply at all sites. We were not able to get reminder reports from the comparison sites because of limited access to those sites (see comment in response to Dr. Mittman), and because not all of the comparison sites activated the reminders. By the time all sites had learned how to generate reports and we had worked out the approvals and other required items, we were in the latter half of the intervention period.*

*Also, it was not possible to obtain historical or retrospective reminder reports. The values of the reminder are over-written in the fields in the clinical database any time a change is made in one of the values in the reminders. This is part of the software, and not within our control. As a result, we were not able to collect any reminder data retrospectively.*

- It is not clear how the self-reported use of reminders was measured. By asking professionals "about their use of reminders generally, and the IHD reminder in particular" is rather vague.

*Our questions on the survey were more specific, and we have included the survey as a supporting file to this revision.*

- page 13, 3rd paragraph: The authors should reflect more on the failure to have a maximum contrast between groups on reminder use.

*We have added a comment to this effect on page 13.*

- The intervention: the reminders were integrated in the Patient Record. "Once triggered they appear in a folder that is available through the face page of the patient's record when it is first opened by the clinician." What does this mean in real practice? Is the clinician always the one who is confronted with the message, or may it be erroneously be the secretarial assistant who prepares the consultations? So, what is the actual exposure?

*Given the passive nature of clinical reminders in the VA, it is not possible to say which member of a provider team views the reminder. However, the actions taken to satisfy or turn off the reminder can only be performed by clinicians authorized to perform the actions. In the case of the measurement reminder, the action required is to order a laboratory test, which in some hospitals can be done by nurses, or can always be done by physicians. In the case of the treatment reminder, only a provider with prescribing authority can make the necessary changes. The "opt-out" options mentioned below require similar authority, and cannot be "turned off" by non-clinical personnel. We also note that in most VHA clinics, there are not many non-clinical personnel, so the likelihood that these reminders would be viewed by a non-clinical provider is remote. However, only clinical personnel with appropriate privileges (detected by their logging in with a unique user name and password, required by VHA privacy and confidentiality policies) are able to perform actions in the reminders.*

- There are escape options for the clinician; one being reasons for deviation of the guideline. This is potentially very interesting information, that might have been integrated in the results as clarification or even a correction factor. The clinician might have very good reasons for deviation, e.g. patient's lipid levels are already known from other sources, or the patient is not willing to start lipid-lowering therapy.

*This kind of information is possible to collect only through the reminder reports. As noted in our previous comment on the reminder reports, we were limited not only in the time period when we were able to collect data, but also in the amount of data we were able to collect. We were only able to obtain data on reminders due and reminders satisfied. We agree that it would be valuable to collect the additional data elements described.*

- In the analysis, data of the beginning and end of the intervention period were compared. Before-after would have been better. It is imaginable that the effect of reminders is largest in the beginning of the intervention period, and this will then weaken the contrast between before and after.

*This is probably correct, although we note that the data at the beginning are actually baseline data—that is, data before the reminders or the intervention were implemented in any site. "After" actually varied by site, with Site C implementing guidelines within a month after the kick-off meeting, Site B sometime within the first 3 months, and Site A approximately six months after the intervention kick-off.*

Minor revisions

- There is quite some unnecessary overlap in text. The writing style is not always clear.

*We have edited the text to reduce any unnecessary overlap and to make the writing clearer.*

- Abstract: "Lowering serum cholesterol has been shown to reduce coronary events, cardiac death and total mortality AMONG HIGH-RISK PATIENTS.

*We have added this important modifier.*

- Abstract: The methods section should contain info on number of participants.

*We have added the number of hospitals in each group and the number of patients at the hospitals during the intervention period.*

- The data are difficult to interpret as they are reported per hospital only in the tables, without summary estimates per trial arm.

*The tables vary. In the first table, much of the information presented is qualitative rather than quantitative, and would be difficult to aggregate to the arms. In Table 2, we present patient level data aggregated only to the arms. We can break this out into hospitals also. In Tables 3a and b, which includes some data available only for intervention hospitals, and some available for both, we have added data comparing the intervention group to the comparison group, as well as data at the hospital level. We In Table 4, we present the aggregate results for the arms first, then broken out by hospital. We divided Table 3 into one table describing reminder use, and another describing survey findings, to make the results clearer.*

- page 9: "We report frequency of reminder use from the reminder reports." What are reminder reports?

*As noted in our response to Dr. Mittman, we have changed the text earlier on to define reminder reports.*

- page 11, footnote. Should you not refer to Table 2?

*Thank you—yes, we have changed this.*

- Page 12: Both performance and attitude data are reported in the first paragraph, which is confusing.

*We have split this paragraph and added a clause to the introductory sentence in the second paragraph to make this less confusing.*

- Women with IHD have been reported in literature to suffer more from under diagnosis and treatment compared to men. Did you have similar findings?

*This is an important topic, but not one we were able to address. The patients treated in the US Veterans Health Administration, as veterans, are typically over 95% male. As a result, we find gender not significant in most of our studies, and we do not report on gender related issues. However, in some sub-analyses done by colleagues using similar data from different hospitals, there is some evidence that women are under-treated. We cannot comment on the issue of under-diagnosis using the data we have used in most of our studies.*

## Ted Palen comments

### Required changes

The p-value reported on page 12 for the 1.77 OR reported for the large intervention site adjusting for self-reported reminder use should be  $<0.05$  NOT  $<0.5$ .

*Thank you for catching this mistake. We have corrected it.*

### Minor comments

1. This reviewer recommends that the footnote on page 7 be incorporated into the body of the reminders section of the Methods. The fact the users of the electronic medical record (EMR) need to go to a separate reminder section of the EMR to see the reminders is a very important point. Reminders of the type described in this intervention are not intrusive and do not require an action to proceed with using the EMR. Therefore, these types are reminders are classified as non-intrusive or passive reminders. These differences are important to highlight in the paper.

*We agree, and have moved the statement about the reminders from the footnote to main text.*

2. The title mentions implementation of the electronic reminders but should also state something about their use, otherwise from the title one thinks this is an implementation study only.

*This is a complex issue, as we have noted in our prior responses. We believe that this paper is most appropriately about implementation of reminders rather than use, given that the focus of our project was on implementing reminders rather than monitoring their use, although if we had been able to collect the reminder reports data as we had planned, we would have had much more robust data about their use. As it is, we do not have these data, and while we can comment about use during one part of the intervention period, we cannot comment throughout the period, nor can we make very strong statements about reminder use. We prefer not to describe the study as reporting on reminder use, given these issues.*

3. The third paragraph in the discussion section is difficult to follow. Rewriting would aid the reader's ability to comprehend the point the authors are making.

*We appreciate this comment. We have deleted the first sentence of this paragraph, which was confusing. We hope that the paragraph now makes more sense than it did.*

4. Comparisons to other studies using other types of alerts (intrusive and non-intrusive) would be helpful in giving insight into which types of alerts prove to work better patient care and clinical workflows.

*We agree, but have found that the literature reporting on reminders and their effect is not very consistent in whether or not the reminders are described in enough detail to know whether they were fully passive or active. We have added a sentence on page 8 to make it clear that the passive nature of VHA reminders was one reason we undertook the intervention effort. We have added a paragraph in the Discussion section (pg. 17) to highlight this issue, and make the point that the passive nature of the VA reminders (which is true of all VA clinical reminders—which are different from VA alerts, which are*

*active, but can be “turned off” quite easily) could well play a role in whether or not they are effective in producing the kind of change desired.*

Martin Lee comments

I have serious concerns about the design of the study and the statistical approaches used. I would like to see responses to these points before consideration of the acceptance of the paper.

*We appreciate the concern, but believe that it relates to a different kind of study than the one we are reporting. We have attempted to clarify our intentions and to make it clear that we were not conducting a trial of the effectiveness of clinical reminders.*

1)pg.5: I am not clear as to why the authors chose not to use cluster randomization (possibly stratified by the size of the institution). Instead they arbitrarily assigned the institutions to the two study arms based on the whether they were in the eastern or western side of the Network. With only 6 institutions, the need for randomization to help in some small way in eliminating or reducing bias is essential. This, however, raises a very fundamental issue with the study. I could not find any indication of a sample size calculation. This points to the concern with the 6 clusters (institutions) and how the authors came up with this number and, in turn, the number of subjects per cluster. What intracluster correlation needed to be assumed to demonstrate the effect size they were interested in should be elucidated. All of this is fundamental to supporting the study design.

*Please see responses to Dr. Mittman and to the other reviews. We have estimated the power based on the sample we had. Because we had no intention of evaluating the effectiveness of reminders, we did not initiate the study with planned accrual or sample size targets.*

2)pg.7: I do not understand the last sentence at the bottom of the page. What is the significance of "...no specific implementation efforts were undertaken"? Please elucidate.

*This is actually central to the results we report, which are related to the intervention to implement the national clinical reminders in the intervention hospitals. Our point is that neither we nor any other group undertook any specific implementation efforts in the comparison hospitals.*

3)pg.9: Why were the number of reminders only collected during the last half of the intervention period? What happened during the first part? Was this an arbitrary delineation to begin collecting data? In this description of data collection, there is no mention of how non-response was handled in order to control for the possible bias that non-response can introduce.

*Please see responses above. There was no “non-response” in the sense of non-compliance with response to surveys in the lack of reminders data—instead, the data were missing from all sites for reasons we have described. Again, we are not making strong statements about evidence regarding the effectiveness of reminders based on the data we present.*

4)pg.10: In the second paragraph, the authors refer to testing "statistical inference". I believe they meant statistical significance. They then go on to state that "only those IHD patients who were present in all time periods during the study period" were included for analysis. This is almost certainly going to introduce bias into the analysis. An intent-to-treat paradigm needs to be incorporated into their analyses. Also on this page, the authors mention the use of cluster

correlation. They need to state how they incorporated this into the analysis, i.e. what methodology was used to adjust the standard errors of their logistic models.

*We have changed the term from “testing statistical inference” to “inference testing” on page 12. We appreciate having this pointed out. We agree that if our intent were to determine the effectiveness of reminders, then using the cohort we did would introduce bias. However, in prior work, we have found that the issue of patients coming into and out of the cohort, throughout an intervention period, also introduces bias. Doing this kind of observational data collection and analysis is quite different from a clinical trial where patients are often completely gone once they exit the cohort. Patients drop in and out of the group of patients meeting the criteria for IHD, but not seeing them in one time period often means only that they did not receive care within VHA during that period. We used several different approaches in this analysis, including analysis by quarters, with and without restricting the cohort. The results we report do not change a great deal. We felt that this approach, given the limited goals we have for this section, and given that it reflects the findings using more complex methods, was reasonable. We have added text on page 13 to describe in more detail how we conducted the multivariable analysis and methods used for cluster correction.*

5)pg.11: The evaluation of the trends in figure 1 are overinterpreted. The fact that the change in guideline compliance is positive and negative over time could simply be random fluctuations. They need to do a time-dependent analysis to determine whether this is or is not the case.

*We appreciate the comment, and have changed the text on page 13 to reduce the over-interpretation. We agree that the trends could be due to random fluctuation, but still believe it is worth presenting the trends as we see them. Given the small number of time points available, we do not believe that time-dependent analysis is justified.*

6)pg.12: In the first paragraph, p-values are needed throughout or at least some assessment of the significance of the comparisons given in that paragraph (confidence intervals, perhaps). The second paragraph discusses the modelling done for each site, or so it appears. I cannot understand why the authors do not do a global analysis comparing the two study arms. Furthermore, ORs are presented, but no confidence intervals. Also, p-values are given as "<0.05". P-values should always be presented as their exact value (to two decimal places).

*We have attempted to clarify the results by giving headings to each section. We have added p-values to the section on provider survey reports of reminder use, and confidence intervals to the section reporting summative patient-level outcomes.*

Basically, I had a hard time trying to decide what was demonstrated here statistically and, therefore, clinically. Their adjustment model appeared to change things, but I am not sure why. I would strongly suggest that the authors revamp their statistical analyses in order for this study to be considered for publication.