Author's response to reviews

Title: How important is clinical decision support in the quality of telephone triage?
A retrospective analysis of triage documentation

Authors:

Frederick North (north.frederick@mayo.edu)
Debra D Richards (richards.debra@mayo.edu)
Kimberly A Bremseth (bremseth.kimberly@mayo.edu)
Mary R Lee (lee.mary2@mayo.edu)
Debra L Cox (cox.debra@mayo.edu)
Prathibha Varkey (varkey.prathibha@emayo.edu)
Robert J Stroebel (stroebel.robert@mayo.edu)

Version: 3
Date: 4 March 2014

Author's response to reviews: see over
We thank the reviewers for their efforts to help us improve this manuscript. Below are our point by point responses to their comments.

**Reviewer 2**

**Comments:**
Overall, the authors have answered most of my points sufficiently, but I have some remarks.

**General:**
I have a question, now I heard that the CDS automatically documents the triage conversation (including all the ticked boxes and outcomes) and suggesting the questions asked. With this knowledge, it seems quite evident that the use of CDS results in improved quality of documentation, at least for part of the criteria used. Some criteria would reach a 100% score with CDS (such as organisation of notes) or a high score (such as number of documented symptoms, which are not necessarily all relevant or improving readability of documentation). In itself, this is still an important message, also with respect to providing information to patients and safety issues. But to me it seems important to be clear about this aspect in the article.

**Author response:** This is a very important point. Documentation (especially computer-generated) does not always improve the readability of the note. We included the following in Discussion to address this.

“Although more documentation is usually good from a risk management standpoint, a potential problem is that it may also be information overload for clinicians who just want a call summary. Documentation of lists of positive and negative symptoms does not necessarily help the readability of the triage note and time-pressed clinicians might dismiss the note altogether. Future study will need to examine how note readability could be improved and how software could highlight the most important positive and negative findings as drivers of triage urgency.”

**Abstract:**
The authors have answered my points sufficiently.

**Reviewer 2:** Perhaps the authors could add the third comparison in the results.

**Author response:** We added the third comparison in the abstract results.

**Background:**

**Reviewer 2:** The authors explained the choice for the extensive background, aiming for two kinds of readers. In that perspective, I can agree with their choice. However, concerning the aim of the study I
am not convinced. The authors give information about telephone triage and the quality of triage, and then ‘jump’ to the aim ‘studying the quality of documentation’. In my opinion it would be helpful to add a few lines, explaining how this aim results from the background information. It is also important to emphasize the difference between quality of documentation and quality of triage, and, as such, these few extra lines would be relevant.

**Author response:**
We agree that quality of documentation not the same as quality of triage. To emphasize this, we added the following to the introduction:  “Quality of triage documentation does not necessarily reflect the quality of triage. For example, a triage nurse could ask an entire set of critical questions to determine the severity of dehydration and then not document it. In this study we solely examine the changes in the quality of triage documentation...”

**Methods:**
The methods are described reasonably well. I have some general comments.

**Reviewer 2:** 1. The additional criteria (number of symptoms and signs, and assessment of overall structure) are criteria that are automatically generated with the CDS. How many of the criteria and AAACN dimensions are automatically registered with the CDS? Some criteria presumably give a 100% score or high score with CDS, due to the automatic registration (which could not be blinded). Perhaps they could made this explicit?

**Author response:**
We made this explicit as requested by the reviewer. In the Appendix notations, for the AAACN documentation measures we put a “+” by the ones that were generated by the CDS and an “*” for the ones that were prompted by the software.

**Reviewer 2:** 1. The nurses in the no-CDS group have some experience with the CDS, but they did not use this for the extracted notes. As they explain this in the methods section and it would probably decrease the potential difference found, this seem to be of minor concern. Yet, it should be mentioned in the limitations section, as they have had used the CDS for a considerable amount of contacts.

**Author response:**
In limitations we added “Also, the group of nurses in the no CDS group may not have been completely comparable to the pre CDS nurses. The findings could be confounded by the selection process of the no CDS nurses who were not as experienced users of ExpertRN and the experience they did have was variable.”

**Reviewer 2:** 2. The authors decided to select the same nurses for the pre-CDS and post-CDS groups, but other nurses for the no-CDS group. Is there a reason for not having a pre-CDS for the other nurses?

**Author response:**
Our object of interest was the CDS intervention. The CDS intervention also contained an element of training so we wanted a concurrent control for a training effect. Having a pre CDS for the concurrent no CDS group would not likely helped us make any better inferences about the impact of the CDS intervention.

**Reviewer 2:** 3. After the revisions, the number ‘22’ nurses got my attention (page 10, selection of triage documents to review). I am not sure if I understand this correctly - are three nurses present in both groups? If so, what could the potential bias be?

**Author response:**
There were three nurses present in both groups. The focus of this study was the CDS intervention. Because we were looking for training effect from nurses trained but not using CDS for those particular notes, it would have been acceptable to use the same group of nurses as long as we could find notes where they didn’t use the CDS. However, having a group of mostly different nurses adds to the credibility of our findings about the CDS intervention. The CDS intervention produces significant differences in documentation across the same nurses after being trained and using the CDS, similar to the differences between them and a mostly different nurse group that was also trained in CDS. We added content to more generally address selection bias in Limitations: “Also, the group of nurses in the no CDS group may not have been completely comparable to the pre CDS nurses. The findings could be confounded by the selection process of the no CDS nurses who were not as experienced users of ExpertRN and the experience they did have was variable.”

**Reviewer 2:** 4. About the ‘care points’ (page 8, in clinical decision support tool): I did not understand this at first as this term is first clarified on page 11. Does the CDS suggest them during the triage contact, or are these separate nurse instructions?

**Author response:**
We added to Methods under Clinical Support Tool (page 8): “These care points are software generated and often offers the nurse a wide range of possible recommendations for patient self-care after the assessment is done.” We also added: “It should be noted that the care points are not just for a home care disposition. For severe chest pain, the software delivers a care point of “chew and swallow one regular-strength aspirin (325 milligrams) or four low-dose aspirin (81 milligrams) as soon as possible...” and “The care points the nurse selects are automatically documented in the triage call record by the software.”

We also added in the appendix that the care points were prompted (indicated by an asterisk).

**Reviewer 2:** 5. The authors added a remark ‘It should be noted that the CDS generated document is not immune to a triagist whose use of the software prompts ...’. I am not sure if I understand this correctly.

**Author response:**
We changed the statement to: “It should be noted that the CDS generated document can be greatly influenced by the choice of the algorithm and path, which is dependent on the critical thinking and skill of the triagist.”

**Reviewer 2:** 6. Perhaps the authors can add the three comparisons in the analysis section (pre-CDS versus CDS, CDS versus no-CDS, no-CDS versus pre-CDS), to clarify these steps for the reader.

**Author response:**
We added that to the analysis section to clarify the analytic steps.

**Results:**

**Reviewer 2:** 1. It appears that the authors used other terminology in table 2 than in the text.

**Author response:**
We thank the reviewer for this observation and we corrected.

**Reviewer 2:** 2. They say ‘there appeared to be no training effect from the extra training associated with CDS’. They actually find a decrease in quality of documentation.

**Author response:**
Thank you for bringing this to our attention. We changed this to “The extra training associated with CDS did not improve the documentation and was associated with a decrease in documentation quality in patient characteristics, contact characteristics, and post-triage disposition.”

**Reviewer 2:** 3. The last paragraph on page 16 is still difficult to read.

**Author response:**
We thank the reviewer for this and have made several changes to improve the paragraph.

**Discussion:**

**Reviewer 2:** 2. Page 18: I would suggest to only mention ‘to document’, as the study did not focus on asking questions. This third paragraph could somehow mislead, as the aim of the study is to describe the quality of documentation. Furthermore, the last sentence in this paragraph also seems to suggest that CDS improves quality of history taking. This seems likely, but should in my opinion be stated with more precautions.

**Author response:**
We changed the sentence so that it states “triage documents” and avoids any inference around history taking: “Using CDS, there were also fewer omissions of critical symptom indicators in the triage documents.”
Reviewer 2: 3. Page 19: Indeed is the documentation of all information helpful for the doctor, but should a doctor rely upon triage history taking? Should a doctor not perform his/her own history taking, also with respect to possible changes in time?

Author response:
This is a very interesting point but we did not feel it was within the scope of our study to discuss. We addressed the readability of the note in this revision, which is something that CDS does contribute to. How much information the doctor should take out of the note delves somewhat into the trustworthiness of telephone triage itself, with or without CDS. How long telephone triage information is useful also gets into the natural history of disease processes and what type of history is taken. The details of an injury might be more accurate from a phone call right after it happens than several days later. On the other hand, a snapshot history of an evolving surgical abdomen would not have much value after a few hours.

Reviewer 2: 4. Page 19: In relation to the first limitation: sometimes nurses seem to document things that were not said in the call (see Derkx et al. ‘Quod scripsi, scripsi.’ The quality of the report of telephone consultations at Dutch out-of-hours centres).

Author response:
We thank the reviewer for bringing this very interesting and important reference to our attention. We added to the limitations that documentation quality does not equal triage quality. We added: “A study by Derkx using incognito patients and comparing recorded calls with triage documents found that the triage documentation at times did not match what was heard on the recorded call. Thus, quality of triage documentation does not necessarily reflect quality of triage.”

This puts the documentation quality aspect of this work into perspective and leads nicely to the Huibers reference which uses robust methodology to evaluate triage quality.

Reviewer 2: 5. Limitation three is not a difference in triage quality but in documentation.

Author response:
We changed “triage quality” to “triage documentation quality” to clarify this.

Reviewer 2: 6. What about the unexpected difference between no-CDS and pre-CDS? Could this also be associated with the selection of groups (bias)? The nurses in the pre-CDS were the same as the nurses in the CDS group, but different from the no CDS group. Furthermore, the authors also mention that the samples had some significant differences across the CDS users and no-CDS control group. What is the potential bias of this?

Author response:
We added the following to explain the possibility of selection bias: “Also, the group of nurses in the no CDS group may not have been completely comparable to the pre CDS nurses. The findings could be
confounded by the selection process of the no CDS nurses who were not as experienced users of ExpertRN and the experience they did have was variable. “

**Reviewer 2:** 7. And, perhaps the authors can start with strengths rather than with limitations.

**Author response:**
We switched the positions and started with strengths.

**Conclusion:**

**Reviewer 2:** In the conclusion I would like to see quality of documentation added to the statement (first sentence

**Author response:**
We changed the first sentence of the conclusion to read: “Telephone triage documentation quality is substantially improved with computerized clinical decision support.” Along with our change in discussion about quality of triage documentation not necessarily reflecting quality of triage, that should remove ambiguity about any conclusion other than for triage documentation.

**Reviewer 3:**
I am content that the changes made reflect the review comments previously made and have improved the paper. HOWEVER, I would still like the fact that there was no a priori sample size calculation made prior to the study (an important element in combating risk to bias in any experimental design) made clear. It should not prevent publication that there was none... but it should be make explicit. Obviously, the confidence intervals imply a degree of precision in any differences found but it is necessary to let the reader know that this is for pragmatic rather than design features (we sampled enough people rather than "we needed X amount of individuals to show a clinically important difference of Y% in the primary measure" and we managed to recruit Z%).

**Author response:**
We explicitly described in Methods that we made no sample size calculations prior to the test sample as follows: “There were no a priori sample size calculations before the test sample.”