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: 2 Date: 29 December 2013

Author’s response to reviews: see over
Author general response on this revision:

The authors greatly appreciate the time and effort of the four reviewers for their helpful comments to improve this manuscript.

Reviewers wanted further statistical information in table 2 including medians, standard deviations, and confidence intervals. We revised table 2 to supply those requested additional statistics.

Reviewers wanted to have an idea about our sample size determination. We included information that this sample size of 50 was determined on the basis of the initial 16 documents reviewed.

Reviewers noted that the study controls needed to be clarified. We added additional information about the triage document controls and additional clarification that the documents were the unit of study and not the triage nurses. We also added specific information about the CDS experience (mean and median uses of ExpertRN) of the nurses who authored the concurrent triage documents but who did not use the assistance of ExpertRN for those specific documents.

Reviewers wanted a defined aim in the abstract. We included our aim in the background section of the abstract.

Responses by reviewer

Reviewer #1: Hiroki Mishina

Reviewer's report:
Thank you for the opportunity to review “How important is clinical decision support in the quality of telephone triage? A retrospective analysis of triage documentation” for consideration for publication in BMC Medical Informatics & Decision Making. The manuscript describes effectiveness of Clinical Decision Support (CDS) in telephone triage.

The manuscript is well written.

I have a concern about the manuscript as currently written:

Reviewer#1
This study design included two cohorts; intervention group (CDS training with CDS tool) and control group (CDS training without CDS tool). Intervention assignment was not randomly. Telephone triage documents were reviewed before and after CDS training. It was not clear to me why authors did not review
telephone triage documents before CDS training from nurses of control group, because I thought that difference between changes of outcome in two groups would show effectiveness of CDS tool. On page 12 and 13, the author wrote “There appeared to be no training effect from the extra training associated with CDS”. I thought authors should evaluate change of outcome after CDS training in control group to determine training effect.

Author response:

This reviewer was not alone in comments concerning the need to improve explanation of study design and cohorts in Methods. To address these comments we extensively edited Methods and added more explanatory detail. In this revision we clarify that the study was conducted at the level of the triage document and we added extra explanation for use of 2 control triage document cohorts. We also added extra explanation of the CDS intervention and the intervention triage document cohort.

Reviewer #2: Linda Huibers

Reviewer's report:

General

Reviewer #2:

The authors describe a study on an interesting and highly relevant research area. Telephone triage has been used in out-of-hours care for decades and is also considered important in day time care. It is a tool to manage the high demands and direct patient flows.

CDS is a relevant topic in this perspective. I would like to hear more about the CDS. Some systems automatically register all symptoms that have been asked in the contact, as the triage professional has to tick boxes in the flowcharts related to the presenting problem. In these systems the numbers of symptoms that are registered are automatically high. Is this the case with the CDS from your study? The aim is to assess the quality of documentation, but if all questions asked are automatically documented I can see that the quality of triage is also assessed (at least in the CDS group). Yet, this is not the case in the other groups, as they have to write notes themselves, which probably influences the level of reporting (especially negative symptoms). Also, if the CDS automatically pops-up questions, the nurses in the CDS group will ask more questions.

Author response: We thank the reviewer for alerting us to make this important clarification. The CDS used in the study, ExpertRN©, does indeed present symptom assessment questions to ask and uses tick boxes to register answers. The CDS also functions to support documentation by automatically generating a note based on the nurses’ interactions with the software. As the reviewer correctly notes elsewhere in her review, the CDS generated document is not immune to a triagist whose use of the
software does not accurately reflect what transpired in the telephone encounter. We added additional description of the CDS in Methods which addresses the reviewer’s comments. We also added more in Discussion about the significance of the documentation support in the CDS.

Reviewer #2:

Abstract Major: 1. Please include the aim of the study. Is it the quality of reporting in a telephone triage contact?

Author response: We added the following aim to the abstract. “The aim of the study was to compare triage documentation quality associated with the use of a clinical decision support tool, ExpertRN©.”

Reviewer #2:

Minor: 2. The results section is quite short and does not give an explicit answer. More symptoms are documented when the CDS tool is used. But are all symptom features relevant? The conclusion also shows some results. Perhaps the authors could add some recommendations.

Author response: We took the result from the conclusion and put it in the result section. We added to the abstract conclusion that “Although this study shows that CDS can improve documentation, further study is needed to determine if it results in improved care.” We also added to Discussion “A more complete evaluation of differences in triage associated with CDS would require a quality review of recorded calls as has been described by Huibers.[22]”

Reviewer #2:

Background Major: 3. The background is quite extensive – the authors should consider shortening it. For example, the third paragraph with examples can be shortened in my opinion, as well as the example on diarrhea and vomiting. 4. The aim of the study does not follow from the background. In the background information is given about telephone triage and the importance of asking the right questions (so called critical symptom indicators) and the insufficient history taking described by some studies in this area. The quality of history taking seems to be the central point of the background - nothing is mentioned on the quality of triage documentation until the authors describe the aim.

Author response: The reviewer makes an excellent point about the background for this study. This manuscript was directed to two audiences, those interested specifically in telephone triage and informaticists who are more generally looking at examples of diverse applications of CDS. For experts in telephone triage, the introduction may be too detailed. However, we wanted to give informaticists sufficient detail to frame why CDS can be important in telephone triage and how it can be applied in this clinical setting.

Reviewer #2:
Minor: 5. Statement of Poole: it would be interesting to read about the exact suggestions.

**Author response:** The Poole reference is a book with a large number of suggestions. We felt that informaticists needed sufficient background to understand the decision making and informatics challenges of triage. The Poole book reference is for those who want more specific suggestions and detailed suggestions outside the scope of the informatics focus.

**Reviewer #2:**

6. The authors state ‘critical symptom indicators’ – can they add a reference? In reference 9 they did not seem to use these critical symptom indicators, but professional clinical judgment.

**Author response:** Critical symptom indicators are indeed based on professional clinical judgment. Originally these were based on consensus opinions by ad hoc groups in the articles referenced. We expanded this critical symptom list for a published manuscript “Should you search the internet for information about your acute symptom?” Telemed J E Health 2012, 18(3):213-218. We mistakenly referenced another article in the original submission and this was corrected (see reference 15). Our method for constructing the critical symptom indicator list is described in full in the reference. Briefly, we reviewed questions from four published adult triage textbooks (Wheeler, Katz, Thompson, Briggs) to develop a set of critical symptom indicators that were common to at least 3 of the 4 textbooks. We referenced our published critical symptom list in Methods for this manuscript. The complete critical symptom indicator list is available in an appendix in our Telemedicine and eHealth paper (reference 15).

**Reviewer #2:**

7. Furthermore, the authors could add more references.

**Author response:** We added to the reference list. Huibers work on triage quality is very relevant to this study, and was added to the references and is included in Discussion.

**Reviewer #2:**

Methods Major: 8. The authors should check headings and clarity. The methods section is quite long – extra headings and paragraphs would help the reader to understand the methods chosen. The methods can be shortened at some place – for example in the paragraph ‘selection of triage documents to review’ there seems to be some duplicate information.

**Author response:** We added an extra subheading for Measures and did extensive editing and explanations as needed to address reviewer concerns.

**Reviewer #2:**
9. Can the authors describe something about their sample size calculation? How did they define the number of 50 per group? (selection of triage documents to review)

**Author response**: We did sample calculations based on the original calibration session. This indicated that a sample size of 50 would be adequate to show significant changes. We included a sentence in Methods to address that: “The 16 record test sample also gave us preliminary information needed to determine that the final sample size of 50 records would be adequate to show significant changes in our major measures.”

**Reviewer #2**:

10. Did the nurses in the control group use the Expert RN? It is stated ‘… but fewer than 300 symptom assessments …’ which implicates that they already used the systems, only not for the notes that were selected. (selection of triage documents to review)

**Author response**: We clarified by restating in Methods that “we used a concurrent control of 50 triage notes from nurses who completed standardized ExpertRN training but had not used ExpertRN for the cohort of notes we evaluated.”

**Reviewer #2**:

11. How did the nurses score the AAACN criteria? (documentation review process and measures)

**Author response**: We clarified this by adding: “The AAACN criteria were scored dichotomously as achieving the criteria or not.” We also added more explanatory material under Statistical Analysis about the scoring (use of majority agreement of 3 nurses).

**Reviewer #2**:

Minor: 12. Please explain what you mean with the following sentence ‘Because telephone triage….., we needed 2 controls’ (Study design and overview)

**Author response**: We further clarified that there were two components of the intervention, the CDS and the nurse who uses it. Because nurses were trained in CDS we needed a control for training only (notes produced by nurses trained in CDS but with notes authored without the CDS). With our design of two controls we juxtapose the CDS-aided note with the note without CDS but authored by the CDS trained nurse. We added the following: “The triagist using the software has to have extra training with the software so it is possible that the extra training alone (no CDS) could improve triage documentation. To address the possible confounding factor of additional training we had a cohort of triage notes authored by nurses with additional training involved in CDS but who did not use CDS for those notes.”

**Reviewer #2**:
13. Already add a short description of the methods, using three measures for assessing the quality of documentation (Study design and overview)

**Author response:** We added a more detailed description of measures just a few paragraphs below with a new subheading Measures.

**Author response:**

14. What is ‘panel’? Is that the number of patients connected/listed to this clinic? (practice setting)

**Author response:** We thank the reviewer for calling to our attention this terminology which is not recognized internationally. We deleted the reference to “panel” and stated the following: “The primary care practice has 141,543 patients. Patients under age 18 account for 15% of the primary care practice and those aged 65 or over are 16% of the total.”

**Reviewer #2:**

15. I am not sure about the meaning of the following sentence ‘Forty-three of the symptom related ... several hundred questions’. Can the choice of one particular symptom at the beginning of triage result in the pop-up of several algorithms, thus leading to several hundred questions? (clinical decision support tool)

**Author response:** We appreciate the reviewer bringing this to our attention and we added additional clarifying text to explain the symptom assessment algorithms and how they are chosen by the nurse. The triage nurses choose from over 100 algorithms which are named by symptom. Of those algorithms some are quite complex and contain several hundred questions which are accessed through branching logic. The branching logic sends the triage nurse to different sets of questions depending on how each question is answered. We added this extra explanatory text under the subheading Clinical Decision Support Tool in Methods.

**Reviewer #2:**

16. The paragraph ‘documentation review process and measures’ is very important for the article, but its readability could be improved. Perhaps the authors could start with mentioning the three measures for assessing the quality of documentation. Also they could consider moving the description of the scoring procedure, with three nurses who used criteria, and a majority rule plus complete consensus. (documentation review process and measures)

**Author response:** We added a Measures subheading with further more detailed explanation of the measures.

**Reviewer #2:**
17. Add the AAACN recommendations as a supplement. (documentation review process and measures)

**Author response:** We put our independently modified criteria that we used in this study as an appendix. Complete AAACN recommendations are available in the reference listed.

**Reviewer #2:**

18. What list was used for assessing the triage notes? Could the authors give more information on this? (documentation review process and measures)

**Author response:** We created the modified criteria (increasing the specificity for reproducibility) that is present in the Appendix. The critical symptom indicators list is contained in the appendix of our referenced Telemedicine and eHealth article (reference 15). Other criteria were counts of content and quality indicators that we explained further in Methods under subheading Measures.

**Reviewer #2:**

Discretionary: 19. The authors could move the sentence ‘Our study design included one CDS intervention group and two control groups’ to the beginning of the first paragraph, after the first sentence. (Study design and overview)

**Author response:** We moved the sentence as suggested.

**Reviewer #2:**

20. The mentioning of comparisons could also be move to the paragraph ‘statistical analysis’, or as a separate paragraph. 21. The authors could consider moving the information about nurse comparison (e.g. level of education) to the next paragraph.

**Author response:** We did extensive editing and rearranging of Methods that address the reviewers’ comments.

**Reviewer #2:**

Results Major: 22. Note disposition is quite different for the ‘No CDS’ group, also concerning the small number of notes in each groups (even though not significant). What are the consequences following the low rate of emergency/911 disposition? The type of contacts in this group seems to be different, with less urgent contacts and follow-up care. The number of questions asked is likely to be lower in these contacts, as an urgent decision is prompted.

**Author response:** We addressed the difference in note disposition in limitations. “Our study design was also retrospective and used real triage notes. We randomly chose notes to review but our samples had some significant differences across the CDS users and the concurrent control.”

**Reviewer #2:**
23. What about the median – does that give approximately the same results? Could the authors add SD to the mean, and 95% CI to the difference? (Table 2)

**Author response**: We added SD and 95% CI to table 2.

**Reviewer #2**:

24. What are ‘... defects in triage note organisation’? Is that the same as ‘major documentation effects’? Where is this information presented – in table 3? How is the 150 possible defects estimated? Can this information be presented in the methods? (paragraph 4 on page 12).

**Author response**: Added the following to methods. “Our triage notes are structured in the electronic medical record so that features of the history are contained in the “History” section of the note, the disposition is contained in the “Impression/Plan” section of the note, and other call demographics are in the header. When reading notes or for software analysis of records, it is important that the information is put in the correct note section. In reviewing the notes, we examined how the information in the notes was organized and if data elements were out of place we classified them as organizational defects. There were 3 sections in each note where data could be misplaced so there was a total of 150 possible organizational defects in the sample of 50.”

**Reviewer #2**:

Minor: 25. Second paragraph presents a lot of information, which is not in a table. Please add ‘not in table’ and try to shorten this.

**Author response**: We were reluctant to truncate this information. The other reviewers did not comment on this and it may be important in their assessment of the study. The important point is the first sentence which states that the nurses were well matched so the reader not interested in details can skim the remainder and the reader looking for rigor can examine the details.

**Reviewer #2**:

26. The authors present % of improvement – I would prefer to also read the exact number.

**Author response**: The exact number is contained in table 2. Manuscript guidelines generally suggest no duplication of table numbers in the text so we decided not to repeat the numbers here. Percent is used here to give an idea of the magnitude without having the reader do the calculation from the table.
27. Perhaps the authors could present an overall score, for all criteria? (Table 2)

**Author response:** We thought an overall score would be presumptuous on our part. It would imply that these criteria were equal (or not equal if we weighted them).

**Reviewer #2:**

28. It would be nice to have more detailed information about the scores per item – are there specific indicators that were problematic? (Table 2)

**Author response:** We added the Appendix containing the criteria used for scoring and also more detailed information in Methods on the scoring.

**Reviewer #2:**

Discussion: Major: 29. Third paragraph page 14: strictly, this study does not say anything about asking questions, but only about registration. In theory, the nurses could write down things that they did not ask in the contact.

**Author response:** In Methods we included “It should be noted that the CDS generated document is not immune to a triagist whose use of the software prompts and checkboxes does not accurately reflect what transpired in the telephone encounter.”

**Reviewer #2:**

30. The subjective assessment by nurses could also be seen as a limitation. Were the nurses blinded, or was it obvious for them which notes were from the post-CDS group, based on the lay-out of the notes?

**Author response:** We added in limitations: “Triage notes could not be blinded to the CDS intervention. The CDS notes were easily identified by the computer generated syntax and the way the note was organized and worded.”

**Reviewer #2:**

31. How can it be that the no CDS groups actually scores worse than the pre CDS group? Can that be related to the number of urgent contacts in the sample? Are there other explanations?

**Author response:** We did a subanalysis of the groups that scored lower and put that into limitations.

**Reviewer #2:**
Minor: 32. Is it correct to state that ‘triage documents authored by nurses trained in CDS but not using it showed no significant improvement’? For this group only one measurement was done, if I understand this correctly?

**Author response:** The same measures were applied to each group of notes and this was clarified in Methods.

**Reviewer #2:**

33. Perhaps the authors could add some headings, such as ‘main findings’, ‘strengths and limitations’, and ‘implications’.

**Author response:** We added a subheading of Limitations and Strengths to set that off from the rest of the discussion.

**Reviewer #2:**

34. The risk management standpoint could be interesting to mention in the background.

**Author response:** We agree that the risk management aspect is very interesting but we decided to leave it up to the readers if they wanted additional information which can be obtained in the Katz reference. We did add to Discussion “this study would be useful for risk managers interested in quantitative evidence of documentation improvement with CDS.”

**Reviewer #2:**

35. Is it indeed better that a doctor already had all the information? Should a doctor not check symptoms and perform his/her own history taking, also in relation to the time factor?

**Author response:** We agree and added that the provider can “focus more on new information or changes to symptoms” to emphasize that the triage documentation can be a point of reference rather than a substitute for collecting new history.

**Reviewer #2:**

36. The first limitation: are telephone contacts not reviewed in malpractice claims? Another study would be to compare the actual conversation with the documentation, as has been done by Derkx et al.

**Author response:** Although many call centers have audio archives of all triage conversations, not all office practices record their calls. For legal or other purposes, the triage document is the only available evidence of what transpired during the call.

**Reviewer #2:**

Tables The authors could improve the readability of the tables, and add more information (as mentioned above).
**Author response:** We added the additional statistics to Table 2 as suggested and improved the headers on the tables.

**Reviewer #2:**

Table 2: 37. Per item there is some improvement, but is it clinically relevant? The mean is presented, I assume – please add this in the legend and table.

**Author response:** In limitations we added that the improvement is in documentation and that it demonstrates that notes showed greater numbers of symptoms assessed but our study design does not allow us to state that this demonstrates clinical relevance. We state that another study looking at differences in clinically relevant outcomes would be a logical follow up to this work.

**Reviewer #2:**

38. Please add what the numbers for ‘triage note content’ mean? Are these absolute numbers, mean?

**Author response:** We added to the table that these are means. In Methods we describe that these are the mean of the counts of the nurse reviewers.

**Reviewer #2:**

Table 3: 39. Quality measures not entirely clear.

**Author response:** We defined the organizational defects and other quality measures more explicitly under a new subheading Measures in Methods.

**Reviewer #2:**

40. For the CDS group there were not defects – is this due to automatic registration?

**Author response:** In Methods we explain that the software automatically transcribes the details of the encounter. This is also discussed in Limitations.

**Reviewer #3:** Carl Thompson

**Reviewer #3:**

Major revisions

state mean level of contacts with CDS in the "no CDS control group" (if its in the order of 1-25 contacts then it really is "noCDS" if its a mean of 200 symptom assessments then it is hardly a "no CDS" control
Author response: We added the mean number of documents for the no CDS control group. We also added to Methods that the CDS not only presented questions and answers it automatically documented them so the intervention had extensive documentation support as well as clinical decision support. We also added this to limitations in Discussion.

Reviewers #3:
measures of variability SD and preferably confidence intervals for mean scores

Author response: We added Median, SD, and confidence intervals to Table 2.

Reviewers #3:

minor essentials
explicit statement of hypotheses
review and discretionary comments

1. Is the question posed by the authors well defined?
There is no research question, hypothesis or explicit aims and objectives posed in the paper. However, the objectives are discernible from the description of the paper in the final paragraph of the background section. An explicit set of hypotheses would help considerably and would aid interpretation (on the part of the reader) of the results if used to frame the presentation of results. E.g. CDS supported nurses will have better quality documentation than unsupported nurses etc. etc.

Author response: We thank the reviewer for pointing this out. We added the explicit aim to the abstract.

Reviewers #3:
2. Are the methods appropriate and well described?
Methods are reasonably well described and the quasi experimental design is appropriate (in the absence of an even sounder randomised and prospective design). The authors have paid attention to the possible impact of training and accommodated this in their design. The (no cds but trained) control group description is a little misleading; as I read this, the nurses could have had up to 299 contacts with a CDS and still have been included in the study as “no CDS” nurses?

Author response: We appreciate the reviewer pointing this out. This was misleading for other reviewers as well. We clarified that the no CDS was at the document level and that the nurses trained in CDS that authored the no CDS notes still had experience with CDS for a significant number of notes. We added further explanation about this and included statistics (mean, median and SD) of the nurse experience with CDS for those who authored the no CDS notes.
Reviewer #3:

Given the common patterns associated with primary care telephone support calls (D&V; sore throats; pain; fever) it is possible that nurses may have had considerable “condition specific” contact with algorithms and the correct questions to ask.

There is no a priori sample size calculation and the measures of central tendency (mean scores table2) have no measures of variation associated with them. At the very least standard deviations are required, but even more preferable would be confidence intervals in order that the reader can judge precision of any effect and the clinical significance of the impacts reported. The lack of reporting of distribution is worrying given the centrality of “variability” in symptom reporting to the background justification for the study.

Author response: In this revision we extensively added statistics including standard deviations, medians and confidence intervals to address the reviewer’s comments (Table 2). We also explained our sample size in Methods.

Reviewer #3:

3. Are the data sound?
   Yes
4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
   APA statistical reporting conventions (if required by the journal) are desirable
5. Are the discussion and conclusions well balanced and adequately supported by the data?
   Yes the authors resist the temptation to go beyond the fact that this was a study examining the effects of CDS on the quality of documentation (not the decisions and judgements that are the focus of the CDS).
6. Are limitations of the work clearly stated?
   Yes, the major limitation is that (legal issues aside) the quality of documentation is not really the purpose of CDS. Rather, it is the quality of choices and judgements – something the authors allude to.

Author response: We completely agree with the reviewer and we added more to Methods and Discussion to distinguish documentation support from decision support.

Reviewer #3:

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
   Yes
8. Do the title and abstract accurately convey what has been found?
Yes although a more declarative title would be preferable (“clinical decision support improves the quality of triage documentation in nurses: a retrospective quasi experiment” perhaps?)

9. Is the writing acceptable?
Yes, its very well written, reasonably concise and to the point.

Author response: We appreciate the helpful comments of the reviewer who has taken time and effort to help us improve this manuscript. In writing this manuscript, we considered a more declarative title as the reviewer suggested but decided on our current title to convey the clinical question that we were seeking to answer. We made the first sentence of the abstract conclusion a declarative summary consistent with the comments of the reviewer.

Reviewer #4: John Chuo

Reviewer #4:
3. Are the data sound?
The data and statistics are persuasive. However, a few lines describing a power calculation (explaining n = 50 in each cohort) would have been helpful. (Minor essential revisions)

Author response: We put in Methods the rationale for the 50 note cohort sample size.

Reviewer #4:
4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
Yes. Of note, it appears that there is a table for each group of measures, except for triage note organization. Could this be added to table 3 if it keeps the manuscript within the size limit? (Discretionary revision)

Author response: We agree with the reviewer that it would be good to fit the organization measure into an existing table. However, since there were 0 to 3 defects, the variable didn’t fit cleanly into the data structure for either Table 2 or Table 3. We could have recoded the defects into a dichotomous variable for each note (organization defects yes/no) which would have matched Table 3 but would have diluted the results.

Reviewer #4:
5. Are the discussion and conclusions well balanced and adequately supported by the data?
Yes. however under limitation 1, the authors argue that ‘the written record is what
counts...however one can argue that from the patient perspective who do not see the written record immediately, the call conversation is what immediately counts. Therefore, glad to see that the authors address need to look at ultimate outcome as the second limitation. It would be helpful to tied limitation 1 with 2 with a transition sentence. (Discretionary revision)

Author response:
We added more to Methods and Discussion about the limitations of this method without having comparison to recorded calls. We revised and added more explanatory material to Limitations that we think address the reviewer’s concern.

6. Are limitations of the work clearly stated? Two limitations are clearly stated. I would add authors should clarify whether the ExpertRN autogenerates any portion of the documentation. I had assumed not, but in the discussion section, the authors mention ‘CDS generated note’. If so, the CDS system, in a sense, has written critical parts of the notes for the nurses. In the overall picture, I am okay with this because the end result is better documentation and it is reasonable to assume that the all the documented points was actually discussed in the conversation; however, this will influence the statistical analysis. i.e. nurses who asked appropriate critical indicator questions but not good documenters are helped more with this system. (Minor essential revision)

Author response: We appreciate the reviewer pointing this out and we added “Our design measures mostly documentation support, and nurses who are poor documenters would be helped more by this CDS.” We also clarified in Methods under the heading Clinical Decision Support tool that the CDS tool,ExpertRN also was, in essence, a documentation support tool. This was also stated this again in Discussion including some more discussion on how risk managers might find this study useful based on quantitative findings of improvement in documentation