Author’s response to reviews

Title: The reliability and accuracy of operational system data in a nationwide helicopter emergency medical services mission database.

Authors:

Anssi Heino (anssi.heino@tyks.fi)
Timo Iirola (timo.iirola@tyks.fi)
Lasse Raatiniemi (lasse.raatiniemi@ppshp.fi)
Jouni Nurmi (jouni.nurmi@hus.fi)
Anna Olkinuora (anna.olkinuora@finnhems.fi)
Päivi Laukkanen-Nevala (pivi@finnhems.fi)
Ilkka Virkkunen (ilkka.virkkunen@finnhems.fi)
Miretta Tommila (miretta.tommila@tyks.fi)

Version: 1 Date: 11 Jul 2019

Author’s response to reviews:

Richard Neville Bradley, M.D.
BMC Emergency Medicine

Reference EMMD-D-19-00033

The reliability and accuracy of operational system data in a nationwide helicopter emergency medical services mission database Anssi Heino, M.D.; Timo Iirola; Lasse Raatiniemi; Jouni Nurmi; Anna Olkinuora; Päivi Laukkanen-Nevala; Ilkka Virkkunen; Miretta Tommila BMC Emergency Medicine
Dear Professor Bradley,

Thank you for your comments concerning our manuscript “The reliability and accuracy of operational system data in a nationwide helicopter emergency medical services mission database”

Please, find below our detailed responses to the reviewer’s comments. The revised manuscript has been uploaded to the Editorial Manager of the BMC Emergency Medicine.

Sincerely,

Anssi Heino

Department of Perioperative Services, Intensive Care Medicine and Pain Management, Turku University Hospital and University of Turku, Turku, Finland

e-mail: anssi.heino@tyks.fi

ASSESSORS' COMMENTS

Referee: 1

Dear authors.

Thanks for having the opportunity to review your manuscript, and congratulations on this interesting work. The study is systematically performed and well written. The presented results represent new information and would be of interest for colleagues in pre-hospital medicine.

Please see my comments below.

Title:

Informative and precise.

Abstract:

Structured overall. The aim is clear. The main content of the paper is well described in few words.
Background:

Lines 48-50 are very relevant information to describe the background for this study. Aim is well described. However, in the lines 53-55 you argue that this study will help you improve quality of data, but without describing in what way a study like yours could do so. I think you should be concrete and give a short example of how this study could really do so. (What kind of concrete measures could be necessary depending on the results of a study like this?). We need studies like this in order to assess the quality of our documentation of HEMS activity and the corresponding potential for research activity.

OUR RESPONSE: The following text has been added in the manuscript (lines 53-56):

“The results will help to improve the quality of data in clinical quality registries in pre-hospital critical care, as they will show the variables most prone to variation and imprecisions in registration, thus allowing to correct them by further instructions, training and data monitoring.”

Methods:

It is my opinion that you should use a guideline from the EQUATOR-network in this study. I believe SQUIRE would be the most appropriate one, as this is a quality improvement study. In the feedback to the reviewers, please explain point by point if all items in the guideline have been addressed (including reasons for not addressing certain items, if applicable).

OUR RESPONSE: SQUIRE has been used as suggested, and please see the point by point explanation, which is at the end of this document.

Setting, Study design and participants, Ethics:

Well described and adequate.

Data collection:

Well described. The fictional missions seem representative for your normal activity. In addition, they seem well selected in order to test different part of your documentation accuracy. The fact that the scenarios were piloted is an important strength for your method.

Lines 107-110: I think you have chosen an adequate focus when selecting variables of interest for your study. I think your comment in this section on the paradox that time variables are so widely used, but yet so little questioned when it comes to their accuracy, is an important one.

Statistical analysis:
Solid and well described.

Results:

Lines 127-128: Were all bases represented (equally)?

OUR RESPONSE: The following text has been added in the manuscript (lines 120-123):

“All six Finnish HEMS bases were represented, and study participant distribution among these was Vantaa with 10 participants (24%), Turku 9 (21%), Tampere 7 (17%), Oulu 4 (10%), Rovaniemi 6 (14%) and Kuopio 6 (14%).”

Lines 139: I don’t understand what kind of data "Dispatcher code" is. Is it a code for which dispatch central dispatched the HEMS-unit? Please explain in the text.

OUR RESPONSE: The following text (explanation) has been added in the manuscript (lines 143-144):

“Dispatcher for HEMS unit can be one of the national dispatch centers, another EMS unit requesting support or the HEMS unit itself attending a mission.”

Tables are ok.

Vital parameters:

How did you define the time for the second measurement of vital parameters? A certain time after the first measurement (when patient was met), or otherwise? Please explain in the text.

OUR RESPONSE: The following text (explanation) has been added in the manuscript (lines 164-166):

“According to the national HEMS CQR guidelines, the time point of the first measurement is the moment the patients has been met, and the secondary parameters are measured after treatment.”

Multi-patient mission documentation:
Lines 173-174. You write that only 23 out of 42 participants registered all four patients. Is this wrong according to your procedures or should they not be registered because you did not transport them? This is relevant to judge the importance of this finding.

OUR RESPONSE: The following text has been added in the manuscript (lines 180-182): “During this study, there was no exact guidelines on multi-patient missions if all patients met by the HEMS unit should be registered in the CQR, or only those that were treated or transported by the HEMS unit.”

Tables:
All tables are easy to read and contribute to a better understanding of the study` s results.

Discussion:
Lines 219-225: Important and well written. I think it is important that you are concrete in terms of what measures could be taken to improve documentation quality (and you are here). Moreover, I would like to read your interpretation of why the inter-rater reliability of the second measurement of vital parameters is worse than that of the first measurement.

OUR RESPONSE: The following text has been added in the manuscript (lines 224-229):

“but it can also be hypothesised that the correct point in the mission to register the first and second vital measurements remains unclear. This could explain especially the poor reliability found for second measurements when compared to first measurements. Instruction to perform a measurement when the patient is met, is an exact time point, whereas there can be major differences in whether the second vital parameter of a patient is measured at the beginning of treatment, at the end of treatment or just before transportation of a patient.”

In lines 232-235 you describe that a possible solution for increasing the accuracy of vital parameters could be electronic data capture. Could this be a possible solution also for the time variables addressed in lines 209-211? (You write that the time variables have low to moderate inter-rater reliability.)

OUR RESPONSE: The following text has been added in the manuscript (lines 238-240):

“The accuracy of vital parameter and time-related variable documentation could benefit from data gathered automatically by monitoring devices, although these devices have their limitations especially for their usage in unconventional pre-hospital setting.”
Conclusion:

Conclusion seems adequate, easy to understand and correct. In addition to routine, intrinsic evaluation of CQRs, my experience is that repeated lectures/discussions for the registering physicians on how to document uniformly is effective.

Other information:

All ok. References seem reasonable to me.

Referee: 2

The reliability and accuracy of operational system data in a nationwide helicopter emergency medical services mission database The quality control of the data is critical for clinical registries that can be used in research on health outcomes. In this sense, the article by Heino A.et al. is a laudable effort to evaluate the reliability of the data and the information provided by your registry.

The methodology used is somewhat contrived. Validating the reliability of the data using fictional clinical scenarios does not avoid a possible bias. As authors write in limitations, the behaviour, the compliance and the accuracy of professionals may not be the same that when they are with a real patient.

OUR RESPONSE: This is an obvious limitation of our study. However, we have described and out-written this in the manuscript (In “Limitations” and “Discussion” paragraphs). Using real-life HEMS missions for this kind of CQR data reliability evaluation would also be challenging, for each mission and patient is different, thus making it impossible to compare registrations of different persons logging the data in the database. In our study setting we have different persons logging in the same data.

In this sense, being a relatively small group, commitment and motivation can play an important role.

OUR RESPONSE: It can be argued, that this is exactly the same with real-life data collection, when personnel concentrate on patient care and not to proper usage of CQR’s, and data collection can even be considered as secondary and irrelevant part of pre-hospital work.
On the other hand, even if the participants are anonymous, as the bases are not, it can be easy to identify the participants.

OUR RESPONSE: The following text has been added in the manuscript (lines 121-123):

“All six Finnish HEMS bases were represented, and study participant distribution among these was Vantaa with 10 participants (24%), Turku 9 (21%), Tampere 7 (17%), Oulu 4 (10%), Rovaniemi 6 (14%) and Kuopio 6 (14%).”

Study participants work in six different bases, which are located in different parts of Finland. Of these six, Oulu had least participants in our study: four, which is 10% of the 42 participants. As described in the manuscript, no personal details were collected and a personal identification number was used instead. This number was not revealed to the investigators. As a result, the investigators did not know, for example, which on-call physicians working at Oulu actually took part in the study.

Leaving aside this basic concern, there are some issues that should be clarified.

Methods

- Table 1 should be placed in results (where it is cited).

OUR RESPONSE: Table 1 removed and placed to “Results” as suggested

- What kind of information was sent by email? It is an important aspect since it can condition the response of the participants

OUR RESPONSE: Following text has been added in the manuscript (lines 96-97):

“E-mailed information, including data collecting period, study protocol and instructions to use of study database”

- The "data collection" section must include a description and classification of the variables collected, as well as the way in which they are collected. For example, are the operational variables automatically included?

OUR RESPONSE: Following text has been added in the manuscript (lines 101-102):
“All of this data was manually logged in the study database by study participants based on study material”

- lines 115-120. The chosen statistical test should be described, even why of your choice, but not to explain why another type of analysis is not used. This explanation must be removed from methods.

OUR RESPONSE: Text removed from the manuscript as suggested, and following text remains (lines 108-117):

“To measure the inter-rater reliability, the per cent agreement and free-marginal multi-rater kappa were calculated. In per cent agreement, the number of equal variables among raters is divided with the number of overall variables that resulted, providing a measure of agreement between raters. Kappa is a form of correlation coefficient, and contrary to percent agreement, it considers a random agreement factor (20). Free-marginal multi-rater kappa was used to study inter-rater reliability in this study setting for its suitability to studies that have free-marginal distributions, namely when raters do not know a priori the quantities of cases that should be distributed into each category (21, 22, 23, 24). Free-marginal multi-rater kappa can take values from 1 to -1. Values from 0 to 1 indicate agreement better than chance, a value of 0 indicates a level of agreement that could have been expected by chance and values from -1 to 0 indicate levels of agreement that are worse than chance.”

Results

The text with the tables is reiterative. It is suggested to summarize the tables, including in one all the operational variables and in another the clinical variables.

OUR RESPONSE: Tables left as they were, based on 1. referee’s comments: “All tables are easy to read and contribute to a better understanding of the study’s results.”, but the tables can be modified if deemed necessary.

Discussion

- lines 183-187 are dispensable. The discussion should start in the next paragraph (line 188)

OUR RESPONSE: Lines were removed and “Discussion” starts in the next paragraph as suggested.
- line 190. Saying "good" is enough, especially when basic problems are later recognized, such as differentiating the reasons why a mission is cancelled (a very basic issue, there is or not patient).

OUR RESPONSE: Changes have been made as suggested (lines 192-193) and following text remains in the manuscript:

“The quality of operational documentation in this CQR is good”

- line 209. It is surprising that the emergency call has a low accuracy. Usually, the entry of a call in a system leaves an automatic trace, very easy to identify. This problem is not well understood. Perhaps, when describing the variables and how they are collected, this and other aspects could be clarified.

OUR RESPONSE: This problem is explained on the next paragraph (lines 215-220): “The registration of time points may seem simple, but again the documentation is based on the interpretation of national guidelines for the usage of CQR, which may vary. Indeed, it is likely that varying personal conceptions of the definitions of specific time points are the primary reason for inaccurate documentation. For example, in cases where another EMS unit asks the HEMS unit to join a mission, it might be unclear whether to register the time of emergency as the call time of the original call to the dispatch centre or the time when the other EMS unit calls”, and in addition the following text has been added in the manuscript (lines 101-103): “This study focused on operational data, including time variables and operational coding. All of this data was manually logged in the study database by study participants based on study material. This imitates real-life HEMS missions and FinnHEMS database.”

- lines 212-218. It is assumed that the definitions of the variables, including the time points, should be part of the documentation sent by email. This argument is not well understood.

OUR RESPONSE: Manuscript modified based on earlier referee comments above; following text added to manuscript:

“E-mailed information, including data collecting period, study protocol and instructions to use of study database, was sent to all the participants three weeks before the data collection began, and the study documents were sent via post to all six bases. The material was sent to bases at the end of 2016 and they were asked to fill in the study database. The data collection period was from 1 December 2016 to 31 January 2017.

This study focused on operational data, including time variables and operational coding. All of this data was manually logged in the study database by study participants based on study
material. This imitates real-life HEMS missions and FinnHEMS database. Operational data such as time variables are often used in pre-hospital studies and quality control, but the accuracy of these variables is rarely questioned. This study focused on the quality and accuracy of this data in the FinnHEMS database.” (lines 96-105) and “The registration of time points may seem simple, but again the documentation is based on the interpretation of national guidelines for the usage of CQR, which may vary. Indeed, it is likely that varying personal conceptions of the definitions of specific time points are the primary reason for inaccurate documentation. For example, in cases where another EMS unit asks the HEMS unit to join a mission, it might be unclear whether to register the time of emergency as the call time of the original call to the dispatch centre or the time when the other EMS unit calls. In addition to exact instructions, proper guidance and continuous quality control of documentation are very important.” (lines 215-221)

- line 219. Usually the most complicated data are those collected in the field, specially clinical variables, unless they have the possibility of being recorded from various devices, (for example, vital signs or ECG measurements performed with defibrillator monitors). Depending on the conditions, an accuracy clinical record can be difficult.

OUR RESPONSE: Following text has been added in the manuscript (lines 238-240):

“The accuracy of vital parameter and time-related variable documentation could benefit from data gathered automatically by monitoring devices, although these devices have their limitations especially for their usage in unconventional pre-hospital setting.”

SQUIRE

Final revision – 4-29-08

• These guidelines provide a framework for reporting formal, planned studies designed to assess the nature and effectiveness of interventions to improve the quality and safety of care.

• It may not be possible to include information about every numbered guideline item in reports of original formal studies, but authors should at least consider every item in writing their reports.

• Although each major section (i.e., Introduction, Methods, Results, and Discussion) of a published original study generally contains some information about the numbered items within that section, information about items from one section (for example, the Introduction) is often also needed in other sections (for example, the Discussion).
Section or Item description

Title and abstract

Did you provide clear and accurate information for finding, indexing, and scanning your paper? YES, SEE LINES 1-33 IN THE MANUSCRIPT

1. Title
   a. Indicates the article concerns the improvement of quality (broadly defined to include the safety, effectiveness, patient-centeredness, timeliness, efficiency, and equity of care) LINES 1-2
   b. States the specific aim of the intervention LINE 1
   c. Specifies the study method used (for example, “A qualitative study,” or “A randomized cluster trial”) EXPLAINED IN THE ABSTRACT, LINES 17-20, THIS TYPE OF STUDY DOES NOT FIT INTO A METHOD THAT COULD BE USED IN A CONCISE TITLE

2. Abstract

Summarizes precisely all key information from various sections of the text using the abstract format of the intended publication YES

Introduction Why did you start? LINE 15

3. Background Knowledge

Provides a brief, non-selective summary of current knowledge of the care problem being addressed, and characteristics of organizations in which it occurs YES, LINES 37-55

4. Local problem

Describes the nature and severity of the specific local problem or system dysfunction that was addressed YES, LINES 51-55

5. Intended improvement
   a. Describes the specific aim (changes/improvements in care processes and
patient outcomes) of the proposed intervention YES, LINES 53-56

b. Specifies who (champions, supporters) and what (events, observations)

triggered the decision to make changes, and why now (timing) YES, LINES 45-50

6. Study question

States precisely the primary improvement-related question and any secondary questions that the study of the intervention was designed to answer YES, LINES 51-52 AND 54-55

Methods

What did you do? YES, LINES 66-70 AND 80-90

7. Ethical issues

Describes ethical aspects of implementing and studying the improvement, such as privacy concerns, protection of participants’ physical well-being, and potential author conflicts of interest, and how ethical concerns were addressed YES, LINES 72-78

8. Setting

Specifies how elements of the local care environment considered most likely to influence change/improvement in the involved site or sites were identified and characterized YES, LINES 58-64, 91-95 AND 101-105

9. Planning the intervention

a. Describes the intervention and its component parts in sufficient detail that others could reproduce it YES, LINES 86-100

b. Indicates main factors that contributed to choice of the specific intervention (for example, analysis of causes of dysfunction; matching relevant improvement experience of others with the local situation) YES, LINES 91-95
c. Outlines initial plans for how the intervention was to be implemented: e.g., what was to be done (initial steps; functions to be accomplished by those steps; how tests of change would be used to modify intervention), and by whom (intended roles, qualifications, and training of staff)
YES, LINES 66-68, 94-95 AND 96-105

10. Planning the study of the intervention

THE TERM “INTERVENTION” DOES NOT QUITE FIT OUR STUDY SETTING, BUT THE WRITTEN HEMS MISSIONS CAN BE CONSIDERED AS AN INTERVENTION AS WE EVALUATED THEIR “RESPONSE” AND “OUTCOME” ON CLINICAL QUALITY REGISTRY

a. Outlines plans for assessing how well the intervention was implemented (dose or intensity of exposure) YES, SEE LINES 81-95

b. Describes mechanisms by which intervention components were expected to cause changes, and plans for testing whether those mechanisms were effective YES, SEE LINES 91-95

c. Identifies the study design (for example, observational, quasi-experimental, experimental) chosen for measuring impact of the intervention on primary and secondary outcomes, if applicable YES, SEE LINES 66-70

d. Explains plans for implementing essential aspects of the chosen study design, as described in publication guidelines for specific designs, if applicable (see, for example, www.equator-network.org) YES, SEE LINES 87-88, THE STUDY SETTING WAS PILOTED

e. Describes aspects of the study design that specifically concerned internal validity (integrity of the data) and external validity (generalizability) YES, SEE LINES 246-250

11. Methods of evaluation

a. Describes instruments and procedures (qualitative, quantitative, or mixed) used to assess a) the effectiveness of implementation, b) the contributions of intervention components and context factors to effectiveness of the intervention, and c) primary and secondary outcomes YES, LINES 58-117

b. Reports efforts to validate and test reliability of assessment instruments YES, LINES 87-88, STUDY WAS PILOTED
c. Explains methods used to assure data quality and adequacy (for example, blinding; repeating measurements and data extraction; training in data collection; collection of sufficient baseline measurements) YES, LINES 67-70 AND 96-105

12. Analysis

a. Provides details of qualitative and quantitative (statistical) methods used to draw inferences from the data YES, LINES 108-117

b. Aligns unit of analysis with level at which the intervention was implemented, if applicable YES, LINES 121-123

c. Specifies degree of variability expected in implementation, change expected in primary outcome (effect size), and ability of study design (including size) to detect such effects YES, LINES 112-114

d. Describes analytic methods used to demonstrate effects of time as a variable (for example, statistical process control) NOT NEEDED IN THIS STUDY SETTING

Results

What did you find? SEE LINES 119-188

13. Outcomes

AGAIN, “INTERVENTION” DOES NOT FIT THIS STUDY SETTING, OUTCOMES EXPLAINED IN THE “RESULTS” PARAGRAPH, SEE LINES 119-188

a) Nature of setting and improvement intervention

i. Characterizes relevant elements of setting or settings (for example, geography, physical resources, organizational culture, history of change efforts), and structures and patterns of care (for example, staffing, leadership) that provided context for the intervention

ii. Explains the actual course of the intervention (for example, sequence of steps, events or phases; type and number of participants at key points), preferably using a time-line diagram or flow chart

iii. Documents degree of success in implementing intervention components
iv. Describes how and why the initial plan evolved, and the most important lessons learned from that evolution, particularly the effects of internal feedback from tests of change (reflexiveness).

b) Changes in processes of care and patient outcomes associated with the intervention

i. Presents data on changes observed in the care delivery process

ii. Presents data on changes observed in measures of patient outcome (for example, morbidity, mortality, function, patient/staff satisfaction, service utilization, cost, care disparities)

iii. Considers benefits, harms, unexpected results, problems, failures

iv. Presents evidence regarding the strength of association between observed changes/improvements and intervention components/context factors

v. Includes summary of missing data for intervention and outcomes

Discussion

What do the findings mean? SEE LINES 190-243

14. Summary SEE LINES 257-261

a. Summarizes the most important successes and difficulties in implementing intervention components, and main changes observed in care delivery and clinical outcomes

b. Highlights the study’s particular strengths

15. Relation to other evidence SEE LINES 48-50

Compares and contrasts study results with relevant findings of others, drawing on broad review of the literature; use of a summary table may be helpful in building on existing evidence

16. Limitations SEE LINES 245-254

a. Considers possible sources of confounding, bias, or imprecision in design, measurement, and analysis that might have affected study outcomes (internal validity) LINES 246-250

b. Explores factors that could affect generalizability (external validity), for example: representativeness of participants; effectiveness of implementation; dose-response effects; features of local care setting SEE LINES 247-248
c. Addresses likelihood that observed gains may weaken over time, and describes plans, if any, for monitoring and maintaining improvement; explicitly states if such planning was not done. SEE LINES 250-252

d. Reviews efforts made to minimize and adjust for study limitations. SEE LINES 252-254

e. Assesses the effect of study limitations on interpretation and application of results. SEE LINES 250-254

17. Interpretation

a. Explores possible reasons for differences between observed and expected outcomes. SEE LINES 191-243

b. Draws inferences consistent with the strength of the data about causal mechanisms and size of observed changes, paying particular attention to components of the intervention and context factors that helped determine the intervention’s effectiveness (or lack thereof), and types of settings in which this intervention is most likely to be effective. SEE LINES 237-243

c. Suggests steps that might be modified to improve future performance. SEE LINES 238-243

d. Reviews issues of opportunity cost and actual financial cost of the intervention. NO COSTS WERE FORMED.

18. Conclusions

a. Considers overall practical usefulness of the intervention. SEE LINES 237-243 AND 256-261

b. Suggests implications of this report for further studies of improvement interventions. SEE LINES 237-243 AND 256-261

Other information

Were other factors relevant to conduct and interpretation of the study? NONE

19. Funding

Describes funding sources, if any, and role of funding organization in design, implementation, interpretation, and publication of study. SEE LINES 269-270