Reviewer’s report

Title: The Impact of EHR and HIE on Reducing Avoidable Admissions: Controlling Main Differential Diagnoses

Version: 2 Date: 6 December 2012

Reviewer: Pedro Pereira Rodrigues

Reviewer’s report:

The manuscript presents an analysis of log-file data in order to inspect whether accessing medical history information from local or external sources can reduce the number of one-day admissions or readmissions within seven days. Data was analyzed for a period of ca. three years (2005-2007) for seven hospitals belonging to the same health maintenance organization. Descriptive statistics of referrals and admissions is presented, along with multiple logistic regression to assess the relevance of patient-related and information-access-related factors on the readmissions and one-day admissions. Results show that access to data (especially external data from other health care institutions) is associated with a decrease in the probability of readmission of one-day admission. However, although this reviewer has found that the topic is extremely important and the research idea is relevant, the implementation is fair, and the analysis (hence, the interpretation) is poor. General comments will elaborate on this.

C1: The manuscript is easy to read in terms of language and verbosity.

C2: The topic is well motivated. There is sound evidence from state-of-the-art literature regarding the benefits of accessing medical history information. However, most of this evidence is lost in descriptions located outside the background section.

C3: Regarding the abstract, background is clear, but the methods would require a bit more detailed exposition.

C4: The research question and hypotheses should appear in the abstract.

C5: Likewise, the results section of the abstract is not precise enough. How was association assessed in the analysis? How does the methodology answer the research question?

C6: Overall, the organization of the manuscript is poor. There's too many methodology-related definitions in the background, and too many background knowledge and references in the methods and results section. The authors should clearly segment the text in the provided sections (background, methods, results, discussion, conclusion).

C7: The separation between research question and research hypotheses is worth noticing. However, it should be more synthetic (no citations to references
for motivation of the hypotheses should appear here; just the hypotheses).

C8: On page 5, you “assumed that some of the single-day admissions were incorrect”. How could you tell these from the correct ones?

C9: The chosen main diagnoses were chosen by the panel of experts; but were they chosen before the study or after inspecting results? Methods section should clarify this.

C10: As previously said, the background was well motivated and, so far, the manuscript got this reviewer hooked to the problem they try to solve. But the method section (and later on the results one too) revealed too many problems. For starters, what is a track log-file analysis? Since you did not explain properly what this his, the reader would expect this to be a standardized analysis, but this reviewer could not find a definition for that study design.

C11: The use of log-files is a relevant tool. The problem with it is also mentioned as they are usually not used. However, there is a bigger problem which is related to the fact that, due to space issues, they are several times turned off by informatics department of health-care institutions. This could have been discussed in the background, along with all the definitions regarding log-files that were erroneously included in the methods section.

C12: The methodology section is really lacking information on the entire study. It is not sufficient to enumerate the dependent and independent variables (better known as outcomes and factors). How data was collected, how participants were selected, what was the unit of analysis (encounters, accesses, patients, physicians, etc.), etc., should appear here.

C13: Regarding readmission within seven days, how do you define “closely related condition”?

C14: Also, the text block “This measure is widely ... during hospitalization [16,22]” should go into the background section.

C15: Regarding single-day admission, the text block “Existing scales ... information [9, 10,12]” should move into background, as well.

C16: Regarding factors, definition of medical history visualization has also background information that should move into previous section (“Vest [23] found...”).

C17: All variables defined here include an explanation of the coding (helpful for logistic regression, later on, but irrelevant here) which is confusing. The authors should define factors and outcomes with proper categories (e.g. readmitted or not, single-day admitted or not, member vs not-member of main HMO, etc.).

C18: The entire subsection of statistical analysis inside the methods section is missing. How did you descriptively analyzed the data? Which measures were used? Which associations were tested? Which statistical tests were thus used? What modeling methods were used? What was the chosen significance level?
C19: Again, there is a mix of methods and results in the results section. Should clearly separate the two types of text.

C20: The initial descriptive statistical analysis is missing. How many accesses? How many patients? How many professionals? How many readmissions? How many single-day admissions? Should expand information contained in Table 2.

C21: Figures 1 and 2 are unnecessary, as the included information is much better understood in a table with proportions. Also, they are not self-contained nor self-sufficient.

C22: The rationale for presenting some results as text and not others is not clear.

C23: Given the specifics of the task at hands, the comprehension of proportions is not always clear (are they related to referrals, admissions, or patients?).

C24: Sorry for the harsh comment, but tables 2-10 are not properly structured. Overall, three-line tables are easier to read.

C25: Regarding table 2: total column should be the last one to the right-hand size (if category labels are presented on the left-hand size).

C26: Regarding table 2: mean plus std dev is difficult to compare (e.g. age); should probably present confidence intervals.

C27: Regarding table 2: percentages should be clearly and easily understood (what is 100% for each case?).

C28: Regarding table 2: length of stay (admission period, in your definition) is rarely a normally distributed variable, often presenting a tail to the right; mean plus std dev are not good summary measures; should probably use median and percentiles.

C29: Regarding table 2: you state that all comparisons are significant under the 0.1% significance level; which comparisons? Between the two groups? This means main HMO and other HMO groups are not comparable regarding the descriptive factors; how does this affect your results?

C30: Authors should, in results, refrain from using qualifiers for found evidence (e.g. “impressive reduction”).

C31: Regarding tables 3-4: There's no descriptive data for the presented groups (e.g. first line of table 3 presents a percentage of readmissions when medical history was not viewed of 4.1%; what is the 100% value for this? It cannot be 3741 stated in table 2, lines 5 or 6). This makes the entire tables unreadable.

C32: Regarding tables 3-4: p-values should always be shown, not hidden in significance levels (unless lower than 0.001, case when stating ‘<0.001’ is acceptable). You should define a significance level (and statistical tests, by the way) in the methods section, and present here the p-values. Significance level should be used in discussion to assess the results.
C33: Regarding tables 5-10: presenting beta values (and standard errors) is redundant as can be computed directly from the OR using log transform. If you prefer to analyze the beta values, you could omit the OR.

C34: Regarding tables 5-10: using boldface to highlight significant results is not consistent; please verify which is the significance level to correspond to a boldface letter.

C35: Overall, it is hard to assess the evidence included in the results, since the confounding factors and issues with the unit of analysis were not properly dealt with.

C36: You state that the work has been approved by the Ethics Committee of the university. What about the HMO or the hospitals where the data belongs? Were patients contacted for an informed consent? Was the data anonymized?

Major Compulsory Revisions
C3-5,C7,C9-10,C12-13,C17-18,C20,C23,C25-29,C31-32,C34-36

Minor Essential Revisions
C2,C6,C11,C14-16,C19,C21,C24,C30,C33

Discretionary Revisions
C1,C8,C22

**Level of interest:** An article of limited interest

**Quality of written English:** Acceptable

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

**Declaration of competing interests:**

I declare that I have no competing interests.