Author's response to reviews

Title: Measuring data reliability for preventive services in Electronic Medical Records

Authors:

Michelle Greiver (mgreiver@rogers.com)
Jan Barnsley (jan.barnsley@utoronto.ca)
Rick Glazier (rick.glazier@ices.on.ca)
Rahim Moineddin (rahim.moineddin@utoronto.ca)
Bart Harvey (bary.harvey@utoronto.ca)

Version: 2 Date: 24 January 2012

Author's response to reviews: see over
Philippa Harris, PhD  
Executive Editor  
BMC-series Journals  
BioMed Central  
Floor 6, 236 Gray’s Inn Road  
London, WC1X 8HL

Dear Dr. Harris,

We appreciate the opportunity to reply to the reviewers’ comments. We have also made changes in the paper to reflect the questions and changes requested. We have included responses to each comment below.

**Reviewer 1**

**Major compulsory revisions**

1. Although a considerable number of patients are involved, the study pertains to only 18 practices. It is an ecological study relating to only this number of practices. It would have been better to link EMRs to the administrative database on a patient by patient basis. In theory, the vaccination rate (for example) in the EMR data, could relate to a totally different set of individuals, as in the administrative database, even though vaccination rates are the same. At least this should be mentioned as a limitation of the study.

Response to 1:

Due to privacy requirements for the Ontario administrative databases, it was not possible to link data at the patient level. The sample size was calculated in order to have a power of 80% to determine a clinically important year over year change of 5% for the provision of each service.[1] Data for several services were aggregated into a composite score, increasing the power of the study. Samples were randomly selected, and there is no a priori reason to believe that the samples differ systematically from the population they were selected from. The administrative dataset reflects the practice population while the EMR dataset reflects samples derived from that population. Nonetheless, we agree that this is a limitation of the study, and this has been added to the paper.

2. The outcomes have been corrected for differences in practice composition using logistic regression. Why was this necessary? The preventive services that are
studied all relate to a specific group within each practice, which implies that correction is not necessary, as the denominator is always comparable across years/practices.

Response to 2:
Our exploratory analyses of factors potentially affecting services in this dataset [2] found that patient age had an effect on the provision of preventive services. We could not be absolutely certain that patient ages were comparable year over year, and therefore used logistic regression to adjust for this variable.

3. I do not understand how the authors come to the conclusion that EMR data for preventive services can not be used to measure and improve quality [of care]. In my view the conclusion should be that the overall difference between medical records and administrative data has decreased with the introduction of electronic medical records (table 4). The fact that the rate between administrative data and electronic medical records was not constant is based is in my view likely to be due to the fact that the study took place shortly after the introduction of EMRs. In my view the conclusion must be that in the first year after the introduction of EMRs these EMRs should better not be used for measuring service provision. Furthermore, reliability seems to vary strongly between the services that are studied.

Response to 3:
We have shown that these data are unreliable in the early stages of EMR implementation. We cannot make conclusions as to reliability of data entry later on, as this was not part of this study. Others [3] have shown that changes in processes for data entry are needed in order to reliably audit preventive processes and drive improvements in quality. It is possible that data entry processes will continue to change over time, and this study provides one method for measuring a domain of data quality. We have added the fact that the study reflects the early phase of EMR implementation.

Minor Essential Revisions

1. A hypothesis is lacking: what do the authors expect to find. Which outcomes would ‘prove’ that EMR data are reliable?

Response to 1:
This has been added in introduction, in the second last paragraph.

Discretionary Revisions

1. The authors sometimes confuse two things: the service having been provided and a service having been recorded and recording of the service (eg. P. 5: the study end point …….; p. 9: we found a lack of improvement in preventive service provision…….

Response to 1:
This has been changed to “documentation”
2. *I do not understand the way paper charts have been examined. If you examine 40 charts per service, you only have a numerator and no denominator.*

**Response to 2:**

We examined 50 charts per physician per service. This is the denominator. The numerator was the number of charts examined with a service being documented.

3. *Tables should include N.*

**Response to 3:**

We have added this information.

4. *p. 8 (top) Table 4 must be Table 5.*

**Response to 4:**

This was corrected in the document.

**Reviewer 2**

**Major Compulsory Revision**

1. *The paper concluded that ‘...we found that data for the preventive services we measured were not reliably entered over time in the medical records we audited when compared to provincial administrative data. Unreliable data cannot be used to measure and improve quality.’ I looked at the data, it seems to me that the greatest differences occurred in the first year after EMR was introduced, which could be explained by ‘teething’ problems in the introduction of a new information system and work processes in clinical practice. In fact Table 4 and Figure 1 suggested that there are signs of convergence of the levels of preventative services recorded by EMR and by Administrative data. The conclusion, particularly the statement ‘data...were not reliably entered over time’ seems rather strong.*

**Response to 1:**

We address this in the reply to Reviewer 1. The text of the conclusion has been changed.

**Discretionary Revisions**

1. *This is more a request for clarification from the authors than a request for revision. The % of those with diabetes (at 8.2%) and mental health (at 21%) seem rather high. Perhaps this is a true reflection of the practice population under study.*

**Response to 1:**
The percentages of patients with diabetes and mental health are derived using validated administrative data algorithms.[4, 5] It is probable that they reflect prevalence in the population studied.

Reviewer 3

Major Compulsory Revisions

1. *I like this paper, but in places I get confused about if you are testing the compliance of best clinical practices related to pay for performance (mammogram, influence, pap), or the data reliability of EMR. This must be clarified throughout the paper. As is, it looks like it is a test of both, and neither is convincing. However, I'm certain it could be strengthened through revision. Also, which components did the EMR have? Clinical decision support? CPOE? Clinical Prompts?*

Response to 1:

This paper addresses data reliability. We looked at compliance with best practices in another paper.[1] We found that EMR implementation did not lead to an increase in the provision of the preventive services we studied, when compared to continued use of paper based records. The EMR used had the ability to generate point of care prompts and reminder letters to patients overdue for a service.

Minor Essential Revisions

1. *Why do you need to use so many administrative data sources?*
   1.1. *Why did all the practices decide on the same system? Are they affiliated?*
   1.2. *The authors should also discuss the future direction of their work.*
   1.3. *Also, what does this mean for practitioners? Policy makers? What are the implications for payers and for researchers?*

Response to 1 (1.1-1.3):

As noted by Reviewer 1, this is an ecological study of 18 physicians. We present data on physician and practice characteristics to provide information on similarities and differences between these physicians and the reader’s setting. The physicians were part of an affiliated group paid through a blended capitation model.[1] They practiced in 9 different office locations. As part of the group contract, they were required to choose a single EMR system. The HITECH act specifies criteria for meaningful use: “EHRs can only deliver their benefits when the information and the EHR are standardized and “structured” in uniform ways, just as ATMs depend on uniformly structured data. Therefore, the “meaningful use” approach requires identification of standards for EHR systems. These are contained in the ONC Standards and Certification regulation announced on July 13, 2010.” (http://www.cms.gov/ehrincentiveprograms/). We have shown that,
early in EMR implementation, data are not reliably entered in medical records; reliability is a prerequisite for data quality. The implication for policy makers, researchers and practitioners is that data quality may be problematic and continued attention to monitoring and improving this area is needed. Standardized and structured data are a requirement for better quality of care.

Thank you once more for giving us the opportunity to reply to the comments.

Michelle Greiver MD CCFP