Sick Patients Have More Data: The Non-Random Completeness of Electronic Health Records

*Nicole G. Weiskopf, MA, *Alex Rusanov, MD, Chunhua Weng, PhD

Columbia University, New York, NY

*Authors with equal contributions.

Abstract

As interest in the reuse of electronic health record (EHR) data for research purposes grows, so too does awareness of the significant data quality problems in these non-traditional datasets. In the past, however, little attention has been paid to whether poor data quality merely introduces noise into EHR-derived datasets, or if there is potential for the creation of spurious signals and bias. In this study we use EHR data to demonstrate a statistically significant relationship between EHR completeness and patient health status, indicating that records with more data are likely to be more representative of sick patients than healthy ones, and therefore may not reflect the broader population found within the EHR.

Introduction

There is great promise in the reuse of electronic health record (EHR) data for clinical research purposes. Retrospective research that reuses existing datasets is generally faster and less costly than prospective research. EHR data have the added benefit of being representative of actual healthcare consumers. For these reasons, amongst others, there is growing interest in the secondary use of EHR data.1-3

When working with EHR data, however, it is important to be mindful of potential data quality caveats.4 A number of studies have established that poor data quality may challenge the suitability of EHR data for research.5-7 In our previous work, we identified three core dimensions of data quality in which researchers engaged in the secondary use of EHR data are interested: completeness, correctness, and currency.8 We further illustrated that EHR data completeness is a task-dependent phenomenon, and that how record completeness is defined depends upon the intended use for the data.9

EHR data completeness may be understood differently in the clinical setting versus in the research setting. Clinically, completeness is most commonly understood to mean fidelity of documentation.6 In other words, a record is considered complete if it contains all information that was observed. When repurposed for secondary use, however, the concept of “fitness for use” can be applied.10,11 In secondary use settings, EHR data completeness becomes extrinsic, and is dependent upon whether or not there are sufficient types and quantities of data to perform a research task of interest. It should be noted that a record deemed incomplete for a given research use might be complete from a clinical standpoint. As an example, the clinical requirement for frequency of blood pressure and heart rate measurement during provision of anesthetic care is every five minutes.12 A record documenting these two variables at five-minute intervals would thus be complete from a clinical standpoint. A researcher, however, might be interested in beat-to-beat variations. A record documenting these variables at five-minute intervals would thus be incomplete from a research standpoint. Similarly, the record would also be incomplete if a researcher was interested in pulmonary artery pressures and these were not documented.

When considering the problem of incomplete data, it is important to understand the manner in which data are incomplete. In statistics, data are understood to be missing at random or missing not at random.13,14 When incompleteness is a random phenomenon, the signal may become noisier, but is otherwise unchanged. Non-random incompleteness, however, may introduce a spurious signal into a dataset. When engaged in the secondary use of EHR data for research, it is important to know whether the records identified as complete for a given research task are representative of all patients of interest, or if the act of selecting for completeness actually results in a non-representative group of records. The literature on this topic is limited.

In the study reported here, we hypothesize the existence of a relationship between record completeness and the underlying health status of the patient of interest. Through the use of representative data types and a broadly applicable definition of complete EHR data, our study proves that the selection of only complete records has the potential to create a biased, non-representative dataset.
Methods

In order to look at the relationship between patient health and record completeness from the secondary use perspective, we required the use of a reliable measure of health status that does not rely entirely on the presence of these data types in the record. The American Society of Anesthesiology (ASA) Physical Classification score is a subjective assessment of overall health or illness severity assigned by an anesthesia provider to every patient requiring anesthetic services. The ASA score is based not only on information present in the EHR but also on interviews with the patient, family and the patient’s other healthcare providers well as information (i.e. laboratory results, imaging and other diagnostic testing results, medical records, prescriptions, etc.) from sources outside the institution and thus not present in the EHR. It is a severity of illness score that is prospectively assigned to each patient by a scoring expert (anesthesia provider) and is thus much less reliant on EHR data than retrospectively assigned scores (e.g. Charlson Comorbidity Index) which rely solely on information present in the EHR. We also sought a measure that is routinely recorded for a broad spectrum of the patient population, as opposed to those that focus on a specific disease (e.g., New York Functional Status Classification for heart failure patients) or segment of the population (e.g., children). The score separates patients into one of six categories (Table 1), with the letter “E” appended for emergency procedures, and has been shown to be strongly correlated with other clinical risk predictors as well as outcomes. Although the requirement of an ASA score limits our cohort to those who have received anesthesia, we felt that it allowed us to capture a more broad population than the other illness measures and have an severity of index measure that was mostly independent of presence of data in the EHR.

The Columbia University Department of Anesthesiology provides anesthetic services for operating rooms at three hospitals – a tertiary-care academic medical center (Milstein), a dedicated children’s hospital (CHONY) and a small community hospital (Allen). Additionally, anesthetic services are also provided for the labor and delivery floors at CHONY and Allen as well as several “off-site” locations such as endoscopy, radiology, neuroradiology, cystoscopy, cardiac catheterization and electrophysiology suites, and ophthalmologic surgery suites. For the majority of patients for whom we provide anesthetic services an anesthesia information management system, CompuRecord (Philips Healthcare, Andover, MA) is used for electronic recordkeeping. The ASA Class score is recorded for all valid procedures. The data collected by this system is periodically migrated to a research database, with the last update occurring in October 2012.

<table>
<thead>
<tr>
<th>ASA Class</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A normal healthy patient</td>
</tr>
<tr>
<td>2</td>
<td>A patient with mild systemic disease</td>
</tr>
<tr>
<td>3</td>
<td>A patient with severe systemic disease</td>
</tr>
<tr>
<td>4</td>
<td>A patient with severe systemic disease that is a constant threat to life</td>
</tr>
<tr>
<td>5</td>
<td>A moribund patient who is not expected to survive without the operation</td>
</tr>
<tr>
<td>6</td>
<td>A declared brain-dead patient whose organs are being removed for donor purposes</td>
</tr>
</tbody>
</table>

The Columbia University Medical Center Institute Review Board approved this study. We began by identifying all patients in the CompuRecord research database with a recorded ICD-9 and CPT code. These codes have been recorded in the database since April 2012, so we identified patients with anesthesia procedures recorded between April, 2012 and the end of September, 2012. To minimize bias introduced by having multiple procedures requiring anesthesia within our time period of interest (one year preceding the procedure), we focused only on the earliest procedure for each patient in our dataset and then excluded all patients who had a procedure in the preceding year, or who were less than one year of age at the time of the procedure. From the remaining set of patients we pulled the ASA Class for a randomly selected set of 5000 patients for the current analysis. Due to the infrequent occurrence of ASA Classes 5 and 6 we dropped the patients in these classes from our analyses.

For demonstration purposes, we selected two common clinical data types through which to assess record completeness: laboratory results and medication orders. Because we did not have a specific use case in mind, we did not use a hard cut-off for defining a complete record. Instead, we assumed that each record existed on a continuum of less or more complete, as implied by counts of days where data were present. The raw counts of days with medication orders and days with laboratory results were calculated for each of the 5,000 patients during the year preceding their procedures using data obtained from the Clinical Data Warehouse (CDW). These data are drawn
from a combination of Allscripts’s Sunrise Clinical Manager for clinical care and Cerner Millennium for ancillary services.

**Data Analysis Methods**

We compared the completeness of records between ASA Classes for both data types using the Kruskal-Wallis one-way analysis of variance. Further post-hoc analyses were performed using the Wilcoxon rank-sum test with a Bonferroni correction.

**Results**

The average age of patients in our sample at the time of surgery was 45 ± 24 years, and 61% were female. The ASA Class distribution for our sample is summarized in Figure 1.

We defined as complete any record having at least N recorded values in each of the two categories (medication orders and laboratory results), where N is the cutoff value, in days. Using this definition, the ratio of complete records over total records in each ASA Class was plotted as N was varied from 1 to 60 days (Figure 2). A magnified view of the plot is shown (Figure 3) to facilitate visualization by improving resolution in the range where N is between 1 and 15. As a specific example, we show the distribution of the proportion of records that are complete across ASA Classes given an arbitrary cut-off of N=7 (Figure 4). As the cutoff of number of desired points increases, the distribution of patients skews further towards the records of patients with an ASA status of 3 or 4.

The average number of days with data for patients in each ASA Class are shown for medication orders and laboratory results in Figure 5. Patients with a more severe health status have a higher number of days with data, on average. A Kruskal-Wallis one-way analysis of variance revealed a significant effect of ASA Class on number of days with medication orders ($\chi^2(3)=332.0, p<0.0001$) and on number of days with laboratory results ($\chi^2(3)=202.2, p<0.0001$). Post-hoc analysis using a Wilcoxon rank sum test with Bonferroni correction showed significant differences between all ASA Classes except Class 1 and Class 2 for both data types, all with p-values <0.0001.

**Discussion**

![Figure 1. ASA Class distribution in study population.](image1)

![Figure 2. Record completeness, per ASA class, over a range of cutoffs (1-60), where cutoffs are the minimum number of values required in each of two categories (medication orders and laboratory results) to make a record complete. For cutoffs >60 all proportions are zero, and thus not shown.](image2)
The results indicate that our hypothesis that there is a relationship between patient health status and EHR completeness is correct. EHR data are not missing at random. Both medication orders and laboratory results show a significant increase in data points per patient as ASA Class increases. Moreover, as the cutoff for what constitutes a complete record increases, the proportion of records considered complete concentrates towards the patients in higher ASA Classes.

This indicates that a sample of patients naively drawn from the EHR—assuming the records chosen are limited to those with sufficient data to be considered complete for the task at hand—will almost invariably be biased towards sicker patients. In other words, those patients with records deemed complete for a given study are unlikely to be representative of the population of interest. These findings are in line with previous work by Collins et al., who found that nursing documentation trends were related to patient mortality. At least in the case of completeness, data quality problems introduce not only noise into EHR data, but bias as well.

A secondary finding suggested by our results is that data consumers may have a difficult time identifying healthy patients with sufficient data for secondary use. This is problematic for those seeking healthy controls or comparison cohorts for research purposes.

**Limitations**

Because of the reliance of our analysis on the presence of ASA Class for all patients in our sample, only anesthesia patients were included in this study. These patients may not be representative of the overall population with data in the CDW. By extension, the CDW itself may not be representative of typical EHR databases due to the tertiary academic medical setting. Nevertheless, we believe that the
essential concept proven by this study—that there is a relationship between record completeness and patient health—will hold true across a broad range of clinical database and institutions.

Future Directions

We intend to continue this research by investigating factors that might be driving the observed trend. For example, we could consider the procedure performed, the diagnosis, the emergency status of the surgery, or inpatient/outpatient status prior to surgery. All are potential confounders.

It would also be helpful to focus on special populations, such as children or pregnant women. One might expect, for example, that pregnant women would be mostly ASA Classes 1 and 2, but with good chart completeness prior to their anesthetic due to pre-natal care. These special populations may require further consideration.

Finally, we hope in the future to identify or develop methodological approaches to selecting complete patient records from EHR databases while avoiding the introduction of bias in the form of patient health status.

Conclusion

In this sample of 5,000 patients the percentage of complete records varied with ASA Class. This relationship held true for laboratory result and medication order data. Using ASA Class as a surrogate for patient health, the data confirm our initial hypothesis that there exists a statistically significant difference between the completeness of electronic health records for sick patients and the completeness of electronic health records for healthy patients. Sicker patients tend to have more complete records and healthier patients tend to have records that are less complete.

These results should serve as a word of caution to researchers wishing to use EHR data for research. Investigators wishing to reuse EHR data must be aware that blind sampling of complete records within an EHR database may skew the sampled population towards sicker patients.

Acknowledgements

We would like to thank Dr. Bruce Levin and Dr. Charles DiMaggio for their assistance in the preparation of this manuscript. This research was funded under NLM grant R01LM009886 and R01LM010815, an NLM training grant 5T15LM007079, and CTSA award UL1 TR000040. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NIH.

References