International Journal of Medical Informatics 111 (2018) 45–50

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

Electronic medical records as a replacement for prospective research data collection in postoperative pain and opioid response studies

Todd Lingren[a], Senthilkumar Sadhasivam[b], Xue Zhang[c], Keith Marsolo[a,*]

[a] Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, USA
[b] Department of Anesthesia, Cincinnati Children's Hospital Medical Center, Cincinnati, USA
[c] Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, USA

ARTICLE INFO

Keywords:
Electronic medical records
Data quality
Dimensional measurement accuracy
Opioid analgesics
Postoperative pain
Respiratory depression

ABSTRACT

Background and aim: Many clinical research studies claim to collect data that are also captured in the electronic medical record (EMR). We evaluate the potential for EMR data to replace prospective research data collection.
Methods: Using a dataset of 358 surgical patients enrolled in a prospective study, we examined the completeness and agreement of EMR and study entries for several variables, including the patient's stay in the post-operative care unit (PACU), surgical pain relief and pain medication side effects.
Results: For all variables with a completeness percentage, values were greater than 96%. For the adverse event variables, we found slight to substantial agreement (Cohen's kappa), ranging from 0.19 (nausea) to 0.48 (respiratory depression) to 0.73 (emesis).
Conclusion: The potential to use EMR data as a replacement for prospective research data collection shows promise, but for now, should be evaluated on a variable-by-variable basis.

1. Introduction

One of the potential benefits of electronic medical records (EMRs) is that capturing data electronically allows the information to be more readily used for purposes other than patient care [1–3]. It remains an open question, however, whether information captured during the course of healthcare operations is truly fit for use in other contexts [4]. Variations in clincial workflows, the ability to document the same information in multiple places in the EMR, and minimal requirements in terms of required fields means that a critical evaluation is necessary before deciding whether to use data captured during the course of clinical care are suitable for research [5,6].

Assessing the quality and completeness of EMR data is a challenge [7–10], and the lack of assessment can cause bias in the results [11]. A recently published data quality framework establishes a number of data quality metrics (e.g., conformance, completeness, plausibility) that can be applied to a given dataset variable as well as the concept of internal and external validation for each (i.e., verification and validation, respectively) [12]. It can be relatively straightforward to establish conformance and completeness measures for a given variable (i.e., is administrative sex always coded as "M" or "F" or "U", and does every patient have a value), but plausibility measures are much more challenging, especially when looking for an external comparator (e.g., does my institution have similar rates of diabetic patients as the general population for given age/sex parameters). At an average academic medical center, hundreds of datasets of EMR data are generated for use in retrospective research every year, yet it is rare for that dataset to be accompanied with a quality scorecard [13], despite such calls from within the informatics community [14]. While one might expect each study team to perform their own quality assessments before running any analyses, it unlikely that they would always have a sufficient understanding of how variations in clinical workflows and documentation practices can impact the downstream quality and availability of a given piece of data.

In this work, we investigate the quality of the documentation relating to post-operative analgesic response in the EMRs of children who underwent tonsillectomies at Cincinnati Children's Hospital Medical Center (CCHMC), between the years of 2009 and 2016. We had a unique opportunity to cross-check the accuracy of the records because these patients were also enrolled in a clinical research study, whereby a member of the research staff was present throughout the child's stay in a post-operative care unit (PACU), collecting clinical outcome data in real-time, independent of the health care providers documenting the patient's vitals, outcomes and clinical status. This study database thus

Abbreviations: EMR, Electronic Medical Record; PACU, Post Anesthesia Care Unit; CCHMC, Cincinnati Children's Hospital Medical Center; MAR, Medication Administration Record; RD, Respiratory Depression; PONV, Post-Operative Nausea or Vomiting; NRS, Numeric Rating Scale; FLACC, Face Legs Activity Cry Consolability

* Corresponding author at: Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave Cincinnati, 45229, USA.
E-mail address: keith.marsolo@cchmc.org (K. Marsolo).

https://doi.org/10.1016/j.ijmedinf.2017.12.014
Received 13 March 2017; Received in revised form 21 September 2017; Accepted 16 December 2017

provides us with an external gold standard for comparison.

Previous work in this area have reported mixed results. Studies have found anesthesia documentation practices to be substandard when compared to established guidelines [15], and when comparing the rates of adverse events, some studies have found manually-reported events to be underreported to those obtained from electronic sources [16], which others have found the opposite [17]. Some of this variation may be explained by the nature of the adverse event and whether it relies on a data stream that is automated (i.e., a feed from a device monitor) versus one that requires manual entry (e.g., vital signs), which speaks to the importance of understanding clinical documentation practices and how they affect whether the underlying dataset actually supports the research question being proposed. Given the paucity of studies that utilize an external comparator, we believe that our ability to compare data extracted from a mature, comprehensive commercial EMR with a prospectively-collected research study database adds to the body of knowledge in this area.

2. Methods

2.1. Dataset description

Five hundred and four children between the ages of 6 and 15 undergoing outpatient adenotonsillectomy with standard perioperative anesthetic, surgical, and nursing care, were enrolled in a clinical study (approved by the CCHMC Institutional Review Board) in which clinical research staff, whose training included daily observations of clinical processes over a one month period, carefully documented the children's post-operative response to analgesics [18,19], including any side effects [19–23]. The PACU nursing staff entered similar data into the EMR (Epic) as part of their normal patient care workflow (we did not consider anesthesia or other inpatient documentation). This involves recording vital signs and other data in EMR "flowsheets" every 15 min, as well as any administered medications, which are documented in the Medication Administration Record (MAR). The information documented in the flowsheets includes the patient's pulse, blood pressure, temperature, respiratory status (including respiratory rate), oxygen saturation level, need for supplemental oxygen and airway support, pain assessments, sedation level, nausea and vomiting, and optionally, any additional notes. We also included the nursing notes associated with the PACU stay as part of our analysis. CCHMC uses a single EMR that is live across all parts of the clinical enterprise. The EMR implementation occurred through a phased roll-out that began in 2007. Each clinic/specialty had the opportunity to customize their EMR configurations by creating fields (typically flowsheets) to allow them to discretely capture the data elements that are needed to support their work in measuring and improving process measures and patient outcomes. The EMR module used by the nursing staff to document information on the patient's stay in the PACU was not live at CCHMC until mid-January 2010, so patients with surgeries prior to that go-live date were excluded, reducing our dataset to 358 patients. The specific variables that we chose to compare are defined and summarized in Table 1.

2.2. Variable description and analysis

For the variables in Table 1, we computed the completeness of each in both the study database and the EMR, defining completeness as whether a value exists or, in the case of a composite variable, whether it can be derived from its relevant subcomponents, for each patient in our dataset. Multiple values for a variable were only counted once. For some variables, particularly post-operative emesis (vomiting), only positive entries are routinely captured. Negative entries may occur, but are less frequent. We see similar patterns with post-operative nausea, though negative entries can be found in the clinical narrative. As a result, we do not report a completeness value for these variables, as it is difficult to determine a true denominator.

To determine a mention of nausea or vomiting in the nursing notes associated with the PACU stay encounter, we searched using regular expressions matching full or partial keywords including (nausea; nauseous; vomiting; vomit; emesis; n/v). Negation was also considered; using a modified version of the NegEx algorithm to exclude false positive mentions [24]. If a patient had positive mentions and negative mentions; the patient was considered to have a net positive mention score. If the patient only had negative mentions; the patient was considered to have a negative mention score. In the emesis-related flowsheets (e.g.; Emesis Occurrence; Emesis Description); any entry that did not indicate a lack of emesis (i.e.; a value of '0' in the Emesis Occurrence field) was considered to be a positive mention.

For total opioid use, we considered the lack of a medication administration record in the EMR to be equivalent to a 0 in the study database (i.e., no opioids given). Since all inpatient medication administrations are captured in the EMR, we are confident that the lack of a record indicates that the event did not occur.

In evaluating a patient's pain, a variety of rating scales were used in the PACU. In our analysis, we focused on Numeric Rating Scale (NRS) scores, as the NRS is more subjective than the others that were used (e.g., FLACC, OUCHER, FACES). There were instances when a patient had an NRS score recorded in the study, but not in the EMR. Because every patient had some form of pain measurement recorded in the EMR, we do not believe that this reflects an incomplete data field. As a result, when computing completeness, we consider the denominator for NRS pain scores to be the number of patients with any pain score recorded in the EMR.

To assess the agreement between the entries in the study database and the EMR for a patient when a value existed in both, we generated Bland-Altman plots for the numeric variables and computed the Cohen kappa coefficient for positive entries of the binary variables (e.g., adverse events) [25].

3. Results

3.1. Completeness

As shown in Table 2, for almost all measurements under consideration, there were comparable degrees of completeness between the EMR and the study database. For every variable where we computed completeness, we found the rate to be > 97%.

3.2. Agreement

Fig. 1 contains the Bland-Altman plots for the numeric variables, illustrating the agreement between the variables as recorded in the EMR and the study database. Fig. 1a shows the time until the post-anesthesia criteria for patient discharge were met [13], and Fig. 1b, the duration of PACU stay. On average, the time until the criteria for discharge were met is 7.75 min more in the EMR than in the study, with 96.1% of the data within the *95% CL* (*95% CL*: −57.17,41.70), and a total of 14 outliers. There was one discrepant instance where the EMR record far exceeded the study record that is out of range on the figure (295 min). For the duration of PACU stay, the mean difference was smaller, at 1.71 min, with 97.2% of the data within the *95% CL* (*95% CL*: −62.12,58.69). There were two discrepant instances with a difference greater than 200 min. One where the study recorded duration of PACU stay far exceeded that available from the EMR (324 min), and another where the EMR record far exceeded the study PACU stay duration (466 min).

Fig. 1c shows the Bland-Altman plot for the total equivalent dosage of the three analgesics (morphine, fentanyl, and hydromorphone) and Fig. 1d, the Bland-Altman plot for the numerical pain scores. The mean difference between the EMR- and study-recorded total dose is 0.012 mg, with 96.3% of the data within the *95% CL* (*95% CL*: −1.12, 1.14). For