Journal of Biomedical Informatics 48 (2014) 66–72

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

Development of a clinician reputation metric to identify appropriate problem-medication pairs in a crowdsourced knowledge base

Allison B. McCoy [a,*], Adam Wright [b], Deevakar Rogith [a], Safa Fathiamini [a], Allison J. Ottenbacher [c], Dean F. Sittig [a]

[a] The University of Texas School of Biomedical Informatics at Houston, 7000 Fannin St., Ste. 600, Houston, TX 70030, USA
[b] Brigham and Women's Hospital, Harvard Medical School, 1620 Tremont St., Boston, MA 02115, USA
[c] The University of Texas Medical School at Houston, 6410 Fannin St., Ste. 1100, Houston, TX 77030, USA

ARTICLE INFO

Article history:
Received 5 September 2013
Accepted 29 November 2013
Available online 7 December 2013

Keywords:
Electronic health records
Crowdsourcing
Knowledge bases
Medical records
Problem-oriented

ABSTRACT

Background: Correlation of data within electronic health records is necessary for implementation of various clinical decision support functions, including patient summarization. A key type of correlation is linking medications to clinical problems; while some databases of problem-medication links are available, they are not robust and depend on problems and medications being encoded in particular terminologies. Crowdsourcing represents one approach to generating robust knowledge bases across a variety of terminologies, but more sophisticated approaches are necessary to improve accuracy and reduce manual data review requirements.

Objective: We sought to develop and evaluate a clinician reputation metric to facilitate the identification of appropriate problem-medication pairs through crowdsourcing without requiring extensive manual review.

Approach: We retrieved medications from our clinical data warehouse that had been prescribed and manually linked to one or more problems by clinicians during e-prescribing between June 1, 2010 and May 31, 2011. We identified measures likely to be associated with the percentage of accurate problem-medication links made by clinicians. Using logistic regression, we created a metric for identifying clinicians who had made greater than or equal to 95% appropriate links. We evaluated the accuracy of the approach by comparing links made by those physicians identified as having appropriate links to a previously manually validated subset of problem-medication pairs.

Results: Of 867 clinicians who asserted a total of 237,748 problem-medication links during the study period, 125 had a reputation metric that predicted the percentage of appropriate links greater than or equal to 95%. These clinicians asserted a total of 2464 linked problem-medication pairs (983 distinct pairs). Compared to a previously validated set of problem-medication pairs, the reputation metric achieved a specificity of 99.5% and marginally improved the sensitivity of previously described knowledge bases.

Conclusion: A reputation metric may be a valuable measure for identifying high quality clinician-entered, crowdsourced data.

1. Introduction

Electronic health records (EHRs) contain vast amounts of data of many types, including medications, laboratory test results, problems, allergies, notes, visits, and health maintenance items. The volume of information is often overwhelming to clinicians and can lead to inefficiencies in patient care [1–4]. Methods for summarizing patient information are required to better organize patient data, which can lead to more effective medical decision making. Developing such summaries requires knowledge about the relationships between the EHR elements [5–7]. Many prior research efforts have described methods for generating this knowledge using standard terminologies [8–10], association-rule mining [11–14], and literature mining [15–17], although each has disadvantages with respect to generalizability, accuracy, and completeness. Crowdsourcing represents a new approach for generating knowledge about relationships between clinical data types that takes advantage of required manual linking by clinicians of these types, such as medications and problems, during e-ordering that

* Corresponding author. Present address: Department of Biostatistics and Bioinformatics, Tulane University School of Public Health and Tropical Medicine, 1440 Canal St., Ste. 2001, New Orleans, LA 70112, USA.
E-mail addresses: amccoy1@tulane.edu (A.B. McCoy), awright5@partners.org (A. Wright), deevakar.rogith@uth.tmc.edu (D. Rogith), safa.fathiamini@uth.tmc.edu (S. Fathiamini), allison.ottenbacher@nih.gov (A.J. Ottenbacher), dean.f.sittig@uth.tmc.edu (D.F. Sittig).

1532-0464/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jbi.2013.11.010

overcomes many limitations of traditional approaches [18]. Initial attempts utilizing this approach showed promise, but there was room for improvement in determining the accuracy of the clinical knowledge [18]. To more accurately classify links, we explored the inclusion of a clinician reputation metric, hypothesizing that such a metric would correlate with the percentage of links made by the clinician that were appropriate.

2. Background

2.1. Clinical summarization

At present, most EHRs present clinical data to providers organized by data type or date [5]. With increasing EHR implementations and growing amounts of patient data, such presentations can hinder point-of-care information retrieval and decision making, leading to clinician dissatisfaction, poor adoption, and substandard patient care [1–5]. Problem-oriented EHRs, or clinical summaries, which organize patient data by relevant clinical problems, make up one approach to overcoming these challenges, but few EHRs have effectively implemented such capabilities [6,7]. One potential cause of low implementation is the limited availability of computable knowledge about the relationships between data elements that is required to develop these summaries.

2.2. Problem-medication knowledge bases

Knowledge bases composed of problem-medication pairs are an important component of clinical summarization. They can also be utilized within EHRs in a variety of other ways, in addition to summarization, such as improving medication reconciliation by grouping together all medications used to treat a particular condition, facilitating order entry by enabling order by indication, and improving the specificity of clinical decision support by enabling different medication dose ranges based on patient condition. However, current procedures for constructing such knowledge bases have significant limitations. The use of standard terminologies or commercially available resources comprises one method, though development of such resources is difficult and expensive, often requiring substantial maintenance [8–10]. Data mining methods are also common but can be hard to execute and may be biased to only include common links [11–13]. Given the drawbacks of these existing methods, new approaches to developing problem-medication knowledge bases are necessary.

2.3. Crowdsourcing

Crowdsourcing is defined as outsourcing a task to a group or community of people [19,20]. This method has been used in various settings to generate large knowledge bases, such as encyclopedias [21]; drug discovery resources [22]; disease treatment, symptom, progression, and outcome data [23,24]; and SNOMED-CT subsets [25]. In recent work, we have applied the crowdsourcing methodology to create a problem-medication knowledge base, which can facilitate the generation of clinical summaries and drive clinical decision support [18]. Fig. 1 depicts an example EHR screen through which clinicians e-prescribe medications (e.g., Aricept 5 MG Oral Tablet) and manually link the medication to the patient's indicating problem (e.g., Alzheimer's Disease). In our crowdsourcing research application, clinician EHR users represent the community, and generating problem-medication pairs for inclusion in the knowledge base represents the task.

Crowdsourcing relies on user input, and the quality of the resulting knowledge depends on correct data collected from the users. In our problem-medication pair application, clinicians may select an incorrect problem for linking due to poor usability, missing problem list entries, or carelessness. As a result, some metrics for evaluating the accuracy of the input for inclusion in a final knowledge base are required. Initial attempts to identify appropriate problem-medication links obtained through crowdsourcing approaches utilized link frequency (i.e., the number of times a problem and medication were manually linked by a provider) and link ratio (i.e., the number of times a co-occurring problem and medication were manually linked by a provider) [18]. However, these measures did not adequately determine the accuracy of all problem-medication pairs, indicating a need for additional metrics for evaluating crowdsourced data.

2.4. Reputation metrics

One method for determining data accuracy utilizes reputation metrics for evaluating user-generated content, such as e-commerce transactions [26], product reviews [27], and e-news or forum comments [28]. Several metrics for evaluating user-generated content have been reported. One approach evaluated feedback on content when a gold standard is not available, generating a reputation metric by comparing an individual's response to others' responses and disseminating ratings to encourage honest, appropriate responses [29]. A later approach expanded these methods, exploring various approaches for identifying true ratings from an aggregated data set [30]. Similarly, an evaluation of product reviews from Amazon.com showed that reviews with a high proportion of helpful votes had a higher impact on sales than those with a low proportion of helpful votes, demonstrating that user-generated content is frequently trusted by other users of a system [31].

More recently, reputation metrics have been applied to evaluating individuals who contribute to crowdsourced knowledge. One group of researchers described reputation and expertise as characteristics of a worker's profile in a taxonomy of quality control in crowdsourcing [32]. In related work, the same authors developed a model for reputation management in crowdsourcing systems; however, like the metrics most frequently described in e-commerce settings, the model requires evaluation of workers by other workers [33]. Another approach used a consensus ratio for evaluating the accuracy of user-submitted map routes, measuring the ratio of agreements and disagreements between users; however, no evaluation of the metric was reported [34]. We hypothesized that these methods could be adapted to evaluate and identify appropriate problem-medication pairs, where clinicians are the users and problem-medication pairs are the user-generated content.

In this study, we developed and validated a clinician reputation metric to evaluate the accuracy of links between medications and problems asserted by clinicians in an EHR during e-prescribing. We hypothesized that the computed reputation metric for a clinician would positively correlate with the appropriateness of the problem-medication pairs that he or she had linked.

3. Methods

3.1. Study setting

We conducted the study at a large, multi-specialty, ambulatory academic practice that provides medical care for adults, adolescents, and children throughout the Houston community. Clinicians utilized Allscripts Enterprise Electronic Health Record (v11.1.7; Chicago, IL) to maintain patient notes and problem lists, order and view results of laboratory tests, and prescribe medications. Clinicians are required to manually link medications to an indication within the patient's clinical problem list for all medications ordered through e-prescribing (Fig. 1). However, medications listed