Journal of Biomedical Informatics 77 (2018) 1–10

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

Learning bundled care opportunities from electronic medical records

You Chen[a,*], Abel N. Kho[b], David Liebovitz[c], Catherine Ivory[d], Sarah Osmundson[e], Jiang Bian[f], Bradley A. Malin[a,g,h]

[a] Dept. of Biomedical Informatics, School of Medicine, Vanderbilt University, Nashville, TN, USA
[b] Institute for Public Health and Medicine, Northwestern University, Chicago, IL, USA
[c] School of Medicine, University of Chicago, Chicago, IL, USA
[d] School of Nursing, Vanderbilt University, Nashville, TN, USA
[e] Dept. of Obstetrics and Gynecology, School of Medicine, Vanderbilt University, Nashville, TN, USA
[f] Dept. of Health Outcomes and Policy, University of Florida, Gainesville, FL, USA
[g] Dept. of Biostatistics, School of Medicine, Vanderbilt University, Nashville, TN, USA
[h] Dept. of Electrical Engineering & Computer Science, School of Engineering, Vanderbilt University, Nashville, TN, USA

ARTICLE INFO

Keywords:
Workflow
Clinical phenotyping
Network analysis
Data mining
Electronic medical record
Topic modeling
Phenotype clusters
Bundled care

ABSTRACT

Objective: The traditional fee-for-service approach to healthcare can lead to the management of a patient's conditions in a siloed manner, inducing various negative consequences. It has been recognized that a bundled approach to healthcare - one that manages a collection of health conditions together - may enable greater efficacy and cost savings. However, it is not always evident which sets of conditions should be managed in a bundled manner. In this study, we investigate if a data-driven approach can automatically learn potential bundles.

Methods: We designed a framework to infer health condition collections (HCCs) based on the similarity of their clinical workflows, according to electronic medical record (EMR) utilization. We evaluated the framework with data from over 16,500 inpatient stays from Northwestern Memorial Hospital in Chicago, Illinois. The plausibility of the inferred HCCs for bundled care was assessed through an online survey of a panel of five experts, whose responses were analyzed via an analysis of variance (ANOVA) at a 95% confidence level. We further assessed the face validity of the HCCs using evidence in the published literature.

Results: The framework inferred four HCCs, indicative of (1) fetal abnormalities, (2) late pregnancies, (3) prostate problems, and (4) chronic diseases, with congestive heart failure featuring prominently. Each HCC was substantiated with evidence in the literature and was deemed plausible for bundled care by the experts at a statistically significant level.

Conclusions: The findings suggest that an automated EMR data-driven framework conducted can provide a basis for discovering bundled care opportunities. Still, translating such findings into actual care management will require further refinement, implementation, and evaluation.

1. Introduction

Under a fee-for-service healthcare model, each of a patient's conditions is managed relatively independently [1,2]. This approach to care can lead to several problems, including delays in (or failure to deliver) service, testing and treatment redundancies, and increased costs for healthcare organizations (HCOs) and patients. In turn, these problems can lead to declines in quality, patient satisfaction, and cost effectiveness [3]. It is anticipated that a shift from fee-for-service to pay-for-value has the potential to resolve, or at least reduce the severity of, many of these problems [4,5]. To realize this alternative vision, HCOs are migrating towards a bundled care model, which is a middle ground between F4S and capitation reimbursement that aims to account for the interplay between various health conditions, rather than focus on each in isolation [6,7].

There are numerous challenges in realizing bundled care. Two of the more pressing are: (1) it is not always evident which health condition collections (HCCs) are appropriate for such a care model and (2) the cost of refining current healthcare systems to support bundled care should be minimized. While HCOs already manage certain complex health needs of patients (e.g., management of comorbidities when treating the primary health problem), such routines often arise in an *ad*

* Corresponding author at: Department of Biomedical Informatics, Vanderbilt University, Nashville, TN 37203, USA.
E-mail address: you.chen@vanderbilt.edu (Y. Chen).

https://doi.org/10.1016/j.jbi.2017.11.014
Received 15 June 2017; Received in revised form 30 October 2017; Accepted 21 November 2017
Available online 22 November 2017
1532-0464/

hoc fashion and are not formalized. As such, there is an opportunity to design a data-driven approach to learn HCCs, which are, or could be, ripe for bundling. A data-driven approach may be further beneficial because, if models are based on the current activities of healthcare professionals, then HCOs could minimize implementation costs for newly established, or the formalization of existing, management routines.

Towards this goal, there is growing evidence that data derived from electronic medical records (EMRs) can be mined to discover associations between health problems [8–14], infer clinical phenomena [15–18], and model healthcare workflow patterns [19–23]. More recently, it has been shown that the relationship between health problems and workflows can be specialized for certain phenomena, such as congestive heart failure [24]. In this paper, we build on such observations and introduce an automated learning framework to discover more general HCCs, which share similar workflows according to the utilization records of an EMR system. We hypothesize that such HCCs could be bundled and managed together based on their shared workflows.

In this paper, we report on how we accomplished this goal by designing a data-driven framework that relies upon a generative topic modeling strategy to infer an association network between HCCs and workflows. We applied a community detection algorithm to infer HCC clusters via the association network. We evaluated this framework with four months of inpatient data (over 16,500 inpatient stays) from Northwestern Memorial Hospital (NMH) and confirmed the plausibility of inferred HCC clusters through an online survey with administrative and clinical experts. We further demonstrated the face validity of HCC clusters through evidence in the published literature.

2. Background

The past several years have witnessed a number of investigations both expert- and data-driven, into modeling and characterizing clinical phenotypes and workflows. However, there has been limited research into automatically establishing relationships between these phenomena. In this section, we review representative expert and data-driven methodologies and illustrate their relationship with our own approach.

2.1. Phenotyping algorithms

Phenotyping aims to use EMR data to detect phenotypes of clinical interest. There has been a flurry of activity in various learning methods for high-throughput phenotyping over the past several years. There are two typical phenotyping algorithms: expert- and data-driven, the latter of which can be further partitioned into supervised and unsupervised techniques.

Expert-driven methods leverage rules developed by experts to identify phenotypes. These methods require substantial manual effort from domain experts. For instance, Kho et. al. developed rule-based phenotyping algorithms to identify subjects for five primary phenotypes (e.g., type 2 diabetes) to support further analysis in genomic association studies [61,62]. Their phenotyping algorithms were developed based on the analysis of EMR data and criteria managed by the Centers for Disease Control and Prevention.

Supervised data-driven phenotyping requires domain experts to exert a substantial of human manual effort to label cases (i.e., phenotypically positive individuals) and controls (i.e., phenotypically negative individuals) and then train and calibrate classifiers. For each new patient record, the classifiers can determine their class label (e.g., case versus control) according to their features. The NIH-sponsored Electronic Medical Records and Genomics (eMERGE) [51] consortium, a network of academic medical centers, has developed a variety of supervised phenotyping algorithms [15,16,52–54]. For instance, Chen et. al. designed a supervised learning algorithm to detect rheumatoid arthritis, colorectal cancer, and venous thromboembolism [52]. Additionally, Zheng et. al. developed a machine learning algorithm to identify type 2 diabetes from EMR data [18].

Unsupervised data-driven phenotyping does not require manual chart review, but instead automatically clusters phenotypes or subtypes [17,24,26,55,56]. The challenge for this type of learning is in the validation of the discovered phenotypes or subtypes. Specifically, there is no clear ground truth for the phenotypes or subtypes that are identified through such an approach. A traditional approach to evaluate the learned phenotypes is to involve clinical and administrative experts to let them review their plausibility [17,64]. For instance, Ho et al. developed a tensor factorization model to automatically identify phenotypes for several major diseases (e.g., metabolic syndrome) from EMR data and, subsequently, recruited expert respondents to review their plausibility. Additionally, Chen et. al. developed a unsupervised learning algorithm to infer phenotypes from EMR data coming from two distinct healthcare systems and validated the similarity, stability and transferability of the learned phenotypes [26].

2.2. Workflow modeling algorithms

Workflow modeling algorithms can be grossly categorized into two types: (i) observational and (ii) data-driven.

Observational studies often rely on manual data collection approaches, such as observations and interviews. One such example was presented by Unertl and colleagues [63], which analyzed direct observations and interviews in hospitals to understand workflow and information flow in the care of chronic diseases. Data-driven algorithms, by contrast, have been proposed to infer clinical pathway patterns through the activity logs of healthcare systems [24,25,57–59]. Almost all of these approaches followed a similar style: (i) infer workflow patterns, then (ii) evaluate the effectiveness of the methods in a clinical case study. For instance, Bouarfa et. al. derived a workflow consensus from clinical activity logs to detect outlying workflows without prior knowledge from experts [57]. They adopted a tree-guided multiple sequence alignment approach to model the consensus of workflows. This strategy was validated over the workflow processes associated with laparoscopic cholecystectomy, where the results indicated the derived consensus conforms to the main steps of the surgical procedure as described in best practice guidelines. In another example, Chen et. al. introduced an altered latent Dirichlet allocation (LDA) based framework to infer clinical workflows through the utilization of an EMR [25] and applied such framework to infer 8 different types of workflows for heart failure patients [24].

To date, phenotyping and workflow modeling algorithms have been developed independently. While there has been little investigation into relating these concepts together, this is vital to the establishment of effective care coordination strategy [60].

3. Research design and methods

The framework for relating phenotypes and workflows is generally composed of four parts: (i) a *workflow inference module*, which is based on the electronically documented actions of EMR users, (ii) a *HCC inference module*, based on patient-specific clinical phenomena indicated in an EMR (e.g., diagnosis codes), (iii) an *association module*, which constructs the association network of HCCs and workflows, and then infer HCC clusters according to the similarity in their workflow patterns and (iv) an *evaluation module*, which consists of online surveys from administrative and clinical experts to determine if the HCC clusters are worthy of consideration for bundling.

We begin with a high-level overview of the models and then proceed with a deeper dive into each component. The general relationships between the workflow module, HCC model and association modeling algorithm are depicted in Fig. 1.

Here, we take a moment to formalize the environment. Let $P = \{p_1,p_2,\cdots,p_n\}$ be the set of patients, $S = \{s_1,s_2,\cdots,s_n\}$ be the set of