ORIGINAL ARTICLE

Performing an Informatics Consult: Methods and Challenges

Alejandro Schuler, MS, Alison Callahan, PhD, Kenneth Jung, PhD, Nigam H. Shah, PhD

Abstract

Our health care system is plagued by missed opportunities, waste, and harm. Data generated in the course of care are often underutilized, scientific insight goes untranslated, and evidence is overlooked. To address these problems, we envisioned a system where aggregate patient data can be used at the bedside to provide practice-based evidence. To create that system, we directly connect practicing physicians to clinical researchers and data scientists through an *informatics consult*. Our team processes and classifies questions posed by clinicians, identifies the appropriate patient data to use, runs the appropriate analyses, and returns an answer, ideally in a 48-hour time window. Here, we discuss the methods that are used for data extraction, processing, and analysis in our consult. We continue to refine our informatics consult service, moving closer to a learning health care system.

Key Words: Learning health system, observational study, practice-based evidence, clinical informatics

J Am Coll Radiol 2017;■:■-■.

INTRODUCTION

Most medical decisions are made without the support of rigorous evidence [1,2] in large part due to the cost and complexity of performing randomized trials [3,4]. Even when guidelines exist, clinicians often do not have the time to read and understand them [1,5]. Furthermore, guidelines often do not apply to complex patients commonly seen in the clinic [1]. In practice, clinicians must use their judgment to make decisions, informed by their own experiences and the collective experience of their colleagues. This often leads to suboptimal care and creates waste and harm [6-8]. Increasingly, it is recognized that the clinical trial enterprise fails to produce relevant evidence for good clinical care [9].

Retrospective observational studies using the electronic health record (EHR) can generate evidence relevant to real patient populations [2]. We have operationalized that opportunity as an *informatics consult* that clinicians solicit the same way they would solicit other specialist consults. Obtaining a consult is a familiar process to clinicians and eliminates the friction between researchers and practitioners, ensuring that practice-based evidence is always readily available. Instead of sending one-way "reports," we offer the consult as a dialogue between the clinician and consult team and among team data scientists so that we are not fooled by oddities in the data and obvious biases. The ultimate goal is to make use of all the evidence on hand to make the best possible decision for patient care.

For example, one clinician requested a consult to assess if the risk of diabetic eye disease is different in diabetic patients treated with rosiglitazone compared with diabetic patients not treated with rosiglitazone. In this case, completing the consult involved an iterative refinement of the analysis to determine an appropriate index time; in the end, we used onset of diabetes as the index time, and after basic matching on age, gender, and length of record, we determined that patients treated with rosiglitazone do not have a statistically significant difference in rate of diabetic eye disease compared with patients not treated with rosiglitazone.

As another example, we received a request from a hospitalist interested in the use of imaging after spinal fusion surgery. The hospitalist requested a consult to determine how many patients who underwent spinal fusion surgery also had a spinal x-ray performed during the inpatient stay when the surgery was performed and in

Center for Biomedical Informatics Research, Stanford University, Stanford, California.

Corresponding author and reprints: Alejandro Schuler, MS, Center for Biomedical Informatics Research, Stanford University, 1265 Welch Road, Stanford, CA 94305; e-mail: aschuler@stanford.edu.

The authors have no conflicts of interest related to the material discussed in this article.

1546-1440/17/$36.00 ■ https://doi.org/10.1016/j.jacr.2017.12.023

the 2 weeks after surgery. We found that the majority of spinal fusion surgery patients had an x-ray taken during their inpatient stay, and fewer than 5% also had a second x-ray taken in the 2 weeks postsurgery.

The generation of good-quality evidence from observational data is not a trivial process, especially when operating on timescales corresponding to the course of care that unfold in days instead of publication schedules that span several months. All observational data are biased in terms of what population is observed (selection bias), what data are recorded on what patients (missing data), and what patients get what treatments (confounding). Depending on the question asked, different methods are required to extract the data, transform it into a useful form, and analyze it to produce evidence [10,11]. In many cases, the methods themselves are being actively researched, and questions remain about their implementation. Naturally, the operational details of the service, which are beyond the scope of this discussion, are equally important as the analysis methods used to generate evidence. We believe that despite these limitations, it is possible to offer a service that uses available data to produce the most up-to-date evidence possible and contextualize the findings for clinicians to incorporate in their decision making.

DATA EXTRACTION AND TRANSFORMATION

Data Sources and Infrastructure

Before beginning the analysis, an appropriate data set must be extracted from the EHR. In our consult, we use data from Stanford's EHR as well as from national claims data sets, such as Truven MarketScan, depending on the question at hand. Our data sets include both structured (eg, International Classification of Diseases, ninth rev codes) and unstructured (text) data. Text data are preprocessed with our text-processing workflow, which has been validated in multiple studies [12,13]. All data elements (e.g. procedures, diagnoses, note text, labs) are mapped to unique clinical concepts using a knowledge graph [14,15]. We anticipate soon having access to linked imaging data, which we will preprocess analogously to text data.

Before proceeding with the consult, we must determine whether we have data that are relevant to the question. We use the Stanford Advanced Temporal Language Aided Search (ATLAS) engine [16,17] to ensure that we have sufficient cohort sizes and data of the required modalities available to complete the consult. The ATLAS engine features a rich temporal query language that enables fast (subsecond response times) and powerful (simple commands define complex logical and temporal restrictions) searches over millions of patient records.

Phenotyping

To perform a search using ATLAS, we must determine the criteria that define the patients of interest (*phenotyping*) [18]. Improper phenotyping can create significant selection biases in the results of downstream analyses [19]. Phenotyping inherently requires domain knowledge because certain criteria may not be clearly or uniquely articulated in EHR data [11]. For instance, to find type 2 diabetic patients, one might search for any patients with an International Classification of Diseases, ninth rev diagnosis code of 250.00, or for patients with 3+ mentions of "t2dm" in their notes, or for patients who are on metformin and have a single mention of "diabetes." Such "rules" to find diabetic patients are often referred to as *phenotyping algorithms*, and it is difficult to judge which is best without expert review [20,21]. We currently rely on the inquiring clinician to help us define an appropriate phenotyping rule for his or her consult.

Supervised machine learning is increasingly used for phenotyping. Instead of defining a rule, a small number of hand-labeled patients are used to train a model, which then classifies the remaining patients [22]. High-specificity rules may also be used to label the training patients [23]. These approaches lessen domain knowledge requirements and may reduce variability in the resulting cohorts. The volume of proxy signals in text and image data make these approaches attractive for labeling phenotypes that are not recorded as structured data in the EHR (eg, socio-economic variables) [14,24]. It may also be possible to include patients in the analysis cohort according to the model's confidence in the phenotype assignment. We are investigating the use of these methods for our consult, but do not currently apply them.

Finally, because phenotype definitions are difficult to evaluate without expert-labeled data, stability analyses are a good way to detect potential biases. If there are multiple alternative phenotype rules or models, the same analysis should be performed using each of them and the final results compared.

Feature Construction

For analysis, each patient must be represented as a vector of *features* that describe their relevant clinical