Journal of Biomedical Informatics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

Using natural language processing to provide personalized learning opportunities from trainee clinical notes

Joshua C. Denny [a,b,*], Anderson Spickard III [a,b], Peter J. Speltz [a], Renee Poirier [c], Donna Rosenstiel [c], James S. Powers [b,c,d,e]

[a] Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN, United States
[b] Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, United States
[c] The Center for Quality Aging, Vanderbilt University School of Medicine, Nashville, TN, United States
[d] The Meharry Consortium Geriatric Education Center, Meharry Medical Center, Nashville, TN, United States
[e] The Tennessee Valley Geriatric Research Education and Clinical Center, Tennessee Valley Healthcare System, Nashville, TN, United States

ARTICLE INFO

Article history:
Received 13 August 2014
Revised 1 June 2015
Accepted 3 June 2015
Available online xxxx

Keywords:
Medical education
Natural language processing
Decision support
Geriatric education
Altered mental status
Advanced directives

ABSTRACT

Objective: Assessment of medical trainee learning through pre-defined competencies is now commonplace in schools of medicine. We describe a novel electronic advisor system using natural language processing (NLP) to identify two geriatric medicine competencies from medical student clinical notes in the electronic medical record: advance directives (AD) and altered mental status (AMS).
Materials and methods: Clinical notes from third year medical students were processed using a general-purpose NLP system to identify biomedical concepts and their section context. The system analyzed these notes for relevance to AD or AMS and generated custom email alerts to students with embedded supplemental learning material customized to their notes. Recall and precision of the two advisors were evaluated by physician review. Students were given pre and post multiple choice question tests broadly covering geriatrics.
Results: Of 102 students approached, 66 students consented and enrolled. The system sent 393 email alerts to 54 students (82%), including 270 for AD and 123 for AMS. Precision was 100% for AD and 93% for AMS. Recall was 69% for AD and 100% for AMS. Students mentioned ADs for 43 patients, with all mentions occurring after first having received an AD reminder. Students accessed educational links 34 times from the 393 email alerts. There was no difference in pre (mean 62%) and post (mean 60%) test scores.
Conclusions: The system effectively identified two educational opportunities using NLP applied to clinical notes and demonstrated a small change in student behavior. Use of electronic advisors such as these may provide a scalable model to assess specific competency elements and deliver educational opportunities.

1. Introduction

Attaining clearly defined competencies is essential for health care professionals. National and international accreditation bodies expect training programs to document the attainment of these competencies by learners [1–4], and with the growing population of elders, more attention has been paid recently to geriatric competency [5,6]. To measure competency-specific performance of medical trainees, educators have focused primarily on standardized milestones and exams, observed structured clinical skills exams, mostly manual logs of clinical exposures, and clinical assessment by attending physicians. Electronic medical records (EMRs) may enable an automated approach that captures students' clinical experiences as a byproduct of their normal clinical work. As a test of a new paradigm for delivering medical education content, we developed an automated education advisor system that analyzed students' EMR notes for relevance to two geriatric competencies and then emailed customized feedback. The two competencies, part of 26 geriatric competencies identified by the American Association of Medical Colleges (AAMC) [5], included: assessment of advanced directives (AD) and evaluation of patients with altered mental status (AMS). We evaluated the accuracy of the system and the effect on each student's knowledge through multiple choice tests.

* Corresponding author at: 400 Eskind Biomedical Library, 2309 Garland Ave, Nashville, TN 37232, United States. Tel.: +1 615 936 5034.
E-mail address: josh.denny@vanderbilt.edu (J.C. Denny).

http://dx.doi.org/10.1016/j.jbi.2015.06.004
1532-0464/

2. Background

2.1. Geriatric competencies for older adults

The AAMC and the John A. Hartford Foundation (JAHF) developed a minimum set of graduating medical student competencies to assure competent care for older patients [5]. Similar efforts have generated competencies for residents and nursing programs (http://www.pogoe.org/geriatrics-competencies). Using the input of leading geriatric educators and survey responses from educators in a number of clinical domains, the consensus panel established eight core geriatric competency domains. Each competency domain contains 2–5 competencies, outlining detailed milestones for medical students in each domain. These competency domains represent an agreed-upon framework to guide curriculum and assessment of medical students. For example, the "cognitive and behavioral disorders" domain includes 5 competencies, which deal with clinical presentation, differential diagnosis, evaluation, and treatment of delirium, dementia, and depression in older adults. These are common diagnoses contributing to AMS presentations in the older adult. Similarly, it is a goal that graduating medical students understand ADs. All hospitalized patients should be assessed for ADs as a result of the Patient Self Determination Act [7]. Given the heightened importance of ADs for elders, we set a clerkship goal that students should discuss ADs with all patients 65 years and older. Competencies for AMS and AD were the target of the present study.

2.2. Competency assessment

Competency-based assessment methods often combine a variety of modalities to provide a comprehensive evaluation of a learner's knowledge and proficiency [8,9]. Medical schools using competency-based assessments typically rely on education portfolios to track each student's progress [10–14]. Portfolio components can include personal reflections on experiences, examinations and their scores, individual and small group projects, simulation encounter reports such as observed structured clinical examinations (OSCEs), mentoring experiences, and clinical exposures. Handwritten log books [15,16] or electronic logs [17,18] allow students to record patient information including demographics, diagnosis, procedures performed, and/or severity of illness. Capture rates of learners' experiences are low because trainees are too busy to enter the data into the system [15,19]. Further, teachers often disagree with students on the primary diagnosis of the case [16,20].

2.3. Overview of the learning portfolio system

To address some of the concerns with manual logs and provide a more robust capture of a trainee's clinical exposure, we developed the KnowledgeMap Learning Portfolio ("Portfolio") system at Vanderbilt, which automatically collects all trainee-authored clinical notes from the EMR as well as providing a forum for other portfolio activities such as personal reflections and specific course content [21]. Medical students are required to write notes (e.g., history and physicals, progress notes, discharge summaries) in the EMR on their patients during their clinical years. From this record, Portfolio creates procedure logs and catalogs patient exposures. Mentors (typically attending and resident physicians) can provide feedback on clinical notes through Portfolio, which has been shown to increase frequency of feedback and improve the quality of the student's assessment and plan in their notes [21]. By applying natural language processing (NLP) through use of the KnowledgeMap concept indexer [22] to identify Unified Medical Language System (UMLS) concepts [23] and SecTag [24,25] to identify note section headers, we have developed search algorithms to automatically map clinical notes to school-identified learning objectives [14,26]. In use since 2005, Portfolio currently includes nearly 5 million indexed trainee-authored notes.

3. Materials and methods

3.1. Setting

We approached third-year medical students at Vanderbilt University School of Medicine during their medicine clerkships between January 2010 through December 2010 with the opportunity to receive learning opportunity messages related to geriatric competencies based on their EMR clinical notes. All students spend at least half of the medicine experience at Vanderbilt University Hospital, the setting for this study; the rest of the patient exposures occur at the Nashville Veterans Affairs hospital whose records were not available to Portfolio. All students in these clerkships were offered the chance to enroll. Students who chose to participate in the study completed a 22-item multiple choice test of geriatric clinical knowledge at the beginning and end of the clerkship. Students enrolled in the study received email alerts for clinical notes matching either of the competencies after completion of the first exam until the completion of their medicine rotation. The Vanderbilt Institutional Review Board approved this study and all participants completed informed consent prior to participation. Course instructors were blinded as to student participation status.

3.2. Education advisor system

We developed a generic framework to allow for creation of electronic education advisors, and then developed specific rules for two geriatric education advisors: altered mental status and advanced directives. These rules were stored in a database and included: a list of students to evaluate; minimum and maximum patient ages to consider; a list of concepts (as UMLS concept unique identifiers) and words to search for within a note; note sections to interrogate; a minimum score threshold required to consider the note a match to the advisor (based on the numbers of concepts and words identified via NLP); boilerplate text for the advisor consisting of "key facts" to send in each email; and whether or not to include a list of relevant medical school curriculum documents relevant to the student's note. The advisor system is flexible, driven by configurable options stored in a database. Search queries and document collections can be built through the KnowledgeMap and Portfolio web interfaces.

Once a student writes a note in the EMR, it was immediately sent from the EMR to Portfolio (as done for all trainee notes as part of the Portfolio system design), which indexed it for all UMLS concepts, tagged with note section according to the SecTag hierarchy of section headers [25]. Information is stored in a relational database. For each advisor, Portfolio analyzed nightly all new trainee notes matching patient characteristics and containing the specified UMLS concepts (see Appendix) located in the specified sections of notes from among all identified UMLS concepts stored in Portfolio (e.g., the core concepts indexed from each note were not restricted to those deemed relevant to the advisors in order to maintain maximum flexibility for other use cases, such as future advisors and searching). Notes exceeding the score threshold resulted in emails to the student with content generated based on the settings, including a link to the relevant note (without personal health identifiers). If the score threshold were set to 0, any note matching the patient characteristics would result in an email, and different text was generated based on whether concepts were or were not

Please cite this article in press as: J.C. Denny et al., Using natural language processing to provide personalized learning opportunities from trainee clinical notes, J Biomed Inform (2015), http://dx.doi.org/10.1016/j.jbi.2015.06.004