Title of Project: Integration of an NLP-based application to support medication management
Principal Investigator: Li Zhou, MD, PhD
Team Members: David W Bates, MD, MSc, Lipika Samal, MD, MPH, Anastasiya Shakurova, PharmD, Qoua L Her, PharmD, MS, Frank Chang, MS, Andrew Seger, PharmD
Organization: Brigham & Women’s Hospital
Inclusive Dates of Project: 7/1/2012-9/30/2014
Federal Project Officer: Steve Bernstein
Acknowledgment of Agency Support: Agency for HealthCare Research and Quality
Grant Award Number: R21HS021544
**Structured Abstract**

**Purpose:** The goal of this study is to develop and assess a natural language processing (NLP) application integrated with an electronic health record (EHR) system to facilitate the medication reconciliation process at the point of care.

**Scope:** The proposed system has been piloted within an ambulatory EHR system at two primary care clinics.

**Methods:** We designed and developed a NLP-based web application, called “NotesLink”, which is built upon a general NLP system (MTERMS) and integrated with an ambulatory EHR system (LMR). The system extracts medication information from clinical notes, applies a knowledge base to compare extracted medications to the patient’s structured medication list, prompts discrepancies alerts (including potentially missing, discontinued, and reactivated medications) to the provider while summarizing and displaying relevant retrieved provider notes in the web-based application. To assess the system performance, we randomly selected two samples of patients who visited the two clinics between December 1, 2013 and May 31, 2014. We first selected 30 patients for a comprehensive evaluation of NotesLink, and then selected 25 patients with at least one medication discrepancy identified by NotesLink to identify and characterize the causes of false positive alerts. The system generated results were compared to the gold standard created by domain experts.

**Results:** Our preliminary findings showed that for the 30 patients, there were 1098 medication mentions (306 unique medications) found in 163 notes. NotesLink achieved a precision of 70.1% and a recall of 73.7% in identifying all medication discrepancies and status. For the 25 patients, 52 medication discrepancies were identified by the NotesLink and 50% of these were possibly true discrepancies according to the manual review.

**Conclusion:** Medication reconciliation process is very complex. Our work identified gaps and challenges of using advanced information technologies combining NLP and automated decision support to facilitate medication reconciliation at the point of care. We also identified new scientific knowledge and made suggestions for future development of comprehensive medication reconciliation systems.

**Key Words:** Natural language processing, terminology, Electronic Health Records (EHRs), medication reconciliation, information retrieval
PURPOSE

Establishing an accurate and complete medication list within a patient’s electronic health record (EHR) is crucial for patient safety [1-12]. Meaningful Use requires healthcare providers to use certified EHRs to perform medication reconciliation to reduce medication errors. To date, numerous efforts have been made to encourage and facilitate the medication reconciliation process at patient care transitions. Clinicians often compare the medical record to an external list of medications obtained from a patient, hospital, or other provider. However, multiple obstacles still exist. First, medication reconciliation in the ambulatory setting is challenging, because clinicians may be unaware of errors due to episodic and often hurried interactions between clinicians and patients with insufficient information exchange [13, 14]. Second, most prior research uses EHR medication lists and pharmacy medication fill histories to identify discrepancies [2, 4, 5, 8, 12]. Certain critical information for medication reconciliation exists in free-text clinical notes that may be unavailable in structured data. In addition, clinicians often need detailed or additional information beyond the medication list in order to make clinical judgments, changes and other decisions. Clinical notes provide an important source of truth. Therefore, innovative tools based on natural language processing (NLP) are needed to retrieve important medication information from notes to facilitate clinician identification of medication discrepancies across different data sources.

SCOPE

Adverse drug events and medication errors are estimated to cost the US health care system $177 billion annually [15]. Medication lists within patients’ EHR are often outdated, incomplete or inaccurate, which is a major cause of medication errors. For example, outdated medications are frequently not deleted. One study [16] shows that 67.4% of medications were still active one calendar day after their inactive status was documented in the clinic notes. Studies have also shown that active medications are often not added to the structured medication list in a timely manner [17-19]. Wagner and Hogan [1] found discrepancies between the number of medications that patients reported taking (5.7) and those listed in their electronic records (4.7). 26% of the discrepancies were related to the failure of the clinician to enter medication changes into EHRs.

Recognizing our vulnerability with regard to medication errors, numerous efforts are underway to encourage health care providers and organizations to perform a medication reconciliation process across the care continuum. The intent is to avoid errors, such as omission, duplication, incorrect doses or timing, and adverse drug-drug, drug-allergy, or drug-disease interactions. Recent efforts strive to establish a reconciled medication regimen through the integration of EHRs with other information sources (e.g., claims, pharmacies, or patients) [2, 4, 5, 8]. Most prior studies have taken place in the inpatient setting. Some electronic inpatient medication reconciliation tools automatically replace the ambulatory EHR medication list with hospital discharge medication orders. This may cause problems because it takes the PCP out of the reconciliation process [11]. There are a few prior studies that address medication reconciliation in outpatient EHRs. Ernst and colleagues [20] found discrepancies in 26.3% of ambulatory charts of patients requesting prescription renewal. Of the charts with discrepancies, 59% omitted medications from the EHR medication list. Other studies [9, 21] found a lack of reporting for whether patients used the medications as originally prescribed, or for cases where their medications were changed by another physician. Lesselroth et al [22] designed ambulatory
check-in kiosks that allow patients to review the drug name, dosage, frequency, and a picture of their medications before their appointment. Medication lists are then retrieved from the EHR and patient updates are captured and reviewed by providers during the clinic session. Schnipper et al [11] designed and developed a tool built into an ambulatory EHR to facilitate post-discharge medication reconciliation.

However, challenges and issues still remain in using health information technology to facilitate the process of gathering, communicating, organizing, and processing medication information across the continuum of care. For example, most current systems are incapable of automated processing and integrating some important data sources (e.g., clinical notes/reports). There is a lack of data harmonization and semantic interoperability between different data sources. In addition, the benefits of medication reconciliation have been difficult to reproduce and implement as reconciling fragmented lists can be labor intensive and time-consuming. Clinicians often need detailed or additional information beyond the medication list (e.g., history and progress of the disease, consultation notes from medical specialists) in order to make judgments, changes and other decisions. There is clear value to using NLP output as a data source for tasks such as medication reconciliation. Using NLP to pull information from textual records and then present that view alongside other data sources, such as the structured medication list in an EHR and prescription fill data in Pharmacy Information Systems, will make these tasks more efficient.

Many NLP tools have been developed for processing biomedical textual data. Details can be found in review articles [23, 24]. Meystre and Haug [25, 26] used NLP to extract potential problem list entries from a list of 80 targeted problems from free-text documents. These problems were then proposed to clinicians for inclusion into a structured problem list. However, to date, very few studies have proposed to use NLP as a complementary means to improving medication reconciliation [16, 27, 28]. Cimino et al [27] combined NLP, a controlled terminology (Medical Entities Dictionary [29]), and a medication classification system (American Hospital Formulary Service Codes [30]) to create metrics to summarize the medication data in both structured and free-text data. Breydo et al [16] developed an algorithm which detects the inactivation of medications in narrative medical documents.

We developed a general NLP system, called the Medical Text Extraction, Reasoning and Mapping System (MTERMS) [31] to extract and encode medication information from electronic clinical notes into a structured format. MTERMS is a modular system using a pipeline approach in which clinical free-text documents are entered into a Preprocessor, to the Semantic Tagger, Terminology Mapper, Context Analyzer, and Parser. The output of MTERMS is a structured document in XML format. MTERMS’ medication lexicon includes a subset of terms from standard terminologies (e.g., RxNorm), local terminologies (e.g. Partners Master Drug Dictionary (MDD), HL7 value sets (e.g. route of administration), regular expression rules, and manually collected terms from chart review or literature review (e.g., misspellings and abbreviations). The Terminology Mapper translates concepts between different terminologies. Partners MDD is used in ambulatory and inpatient EHR systems by providers at the time of ordering. MDD is mapped to First DataBank (FDB)’s ingredient codes (HIC_SEQNO), the Generic Code Number Sequence Numbers (GCN-SEQNO) and the Enhanced Therapeutic Classification (ETC) codes. RxNorm, created and maintained by the National Library of Medicine (NLM), aims to provide a standardized nomenclature that relates itself to terms from commonly used source vocabularies [32]. MTERMS is able to encode medications in free-text
notes using MDD, RxNorm and FDB codes and also establish mappings among them. The Context Analyzer looks for contextual information (e.g., temporal information, negation) to further determine the meaning of a phrase in context with the rest of the text. The goal of this study is to develop novel methods and a tool that will use MTERMS’ NLP output to facilitate the medication reconciliation process in the outpatient setting [34, 35].

METHODS

Overview

In this two-year study we have developed and evaluated a systematic, novel approach to integrate an NLP-based tool to support medication list management at the point of care. Figure 1 shows the study design overview. First, we conducted a qualitative analysis to understand user requirements, use cases, system functional specifications, workflow issues, barriers to and facilitators of using clinical notes in the medication reconciliation process in the ambulatory setting (Aim 1). This information was used to design, develop, and implement a real-time web-based tool, NotesLink, that consists of several novel system components and unique features to facilitate medication reconciliation. NotesLink is a web application built upon the MTERMS NLP system and it has been integrated with our ambulatory EHR system, the Longitudinal Medical Records (LMR). It presents NLP output and links to original notes in an efficient way to facilitate clinicians to review medication discrepancies between notes and the structured medication list to make further necessary changes (Aim 2). We implemented a pilot study at two Brigham and Women’s Primary Care (BWPC) practices, and measured the performance and feasibility of the methods and the NotesLink system in improving medication reconciliation in the outpatient setting (Aim 3). Finally, we presented the results and shared the system design in biomedical informatics conferences and various academic seminars. We are currently preparing several manuscripts targeted for clinical and informatics journals (Aim 4).

Figure 1. Study Design Overview

Requirement Analysis

Since there are no published studies on how to use NLP output to support medication reconciliation and little is known about the barriers to and facilitators of using this approach in the outpatient clinics, it is important to conduct qualitative research to understand the requirements, use cases, workflow and other relevant issues before designing, developing and implementing a computer application. We conducted Focus Group Discussions (FGDs) [36] with two types of subjects: 1) clinicians and 2) technical professionals. FGDs are a widely used qualitative research technique in which a group of people are asked about their opinions, beliefs and attitudes towards a product, service, or idea. FGDs differ from individual interviews in that the discussion allows for interaction among all the members of the group. A trained moderator facilitated the discussions and organized conversation around a set of pre-determined topics detailed in a moderator's guide [37].
Clinician FGDs. The overall goal of the clinician FGDs was to understand user requirements and workflow issues of using the proposed approach for medication reconciliation. The FGD sessions helped us to solicit clinicians’ suggestions on the user interface design. Focus group sessions were held in each of the two clinics. In total, 24 clinicians (including 16 physicians, 3 pharmacists, and 5 nurses) participated in the discussions. We began with a series of warm-up questions to explore clinicians’ experience and opinions about the current medication reconciliation process. Then, we asked the clinicians to describe their workflow for medication reconciliation activities and, particularly, to articulate the ease or difficulty involved in carrying out necessary steps. Investigators shared with the clinicians the proposed NLP-based approach and inquired their input on how to make this tool useful, helpful and efficient.

Technical Professionals FGDs. The goal of the technical professionals FGD session was to understand the current system architecture, the requirements for integrating the proposed tool with the LMR and its current medication reconciliation application, the lessons learned from previous projects, and other technical issues. Eight technical professionals, including developers, analysts and informaticians who have been involved in previous development of medication reconciliation applications, participated in the discussions.

All discussions in FGD sessions were audio-taped for later transcription and analysis. We performed a content analysis of the transcribed FGDs and apply the immersion-crystallization method [38]. These qualitative findings were used to enrich and modify the original NLP system and guided our design and development of the proposed system.

Design, Development and Implementation of the NotesLink System

One principle of the system design is to use a modular architecture, so the application will be portable and can be integrated with other clinical information systems in the future. Therefore, instead of embedding this application into the LMR, we applied Service Oriented Architecture (SOA) by providing an NLP service and other necessary generic services and system functions. The system architecture is shown in Figure 2.

Figure 2. System Architecture (PHS: Partners Healthcare Systems)

Figure 3 shows system components of NotesLink. The Batch Processing Mode obtains each patient’s notes from LMR before their scheduled visit at the clinic. Based on the focus group discussions with clinicians, NotesLink retrieves three years worth of notes per patient before their schedule visit at the clinic. MTERMS performs NLP to extract and encode medication
information from each note, and store the structured XML output into a SQL server database. NotesLink’s *Summarization Mode* obtains the patient’s medications from LMR’s structured medication list, after which it compares them with the medications extracted from notes, and then saves the comparison results back into the database for NotesLink’s *Web Application* and relevant pages of LMR.

Figure 3. System Components

We added an *http* link in LMR that users can launch NotesLink Web application by clicking this link (see Figure 4). This link was placed in LMR’s Patient Summary Page, Medications Page and Medication Reconciliation Page. The link displays two types of information depending on whether NotesLink identified medication discrepancies or not. If discrepancies are found, the link signifies that medications are missing from the medication list or should be discontinued. It also quantifies the number of discrepancies (e.g., “Medication Discrepancies from Notes (2)”). When no discrepancies are found, the link displays a message such as “Medication from Notes”. After the user clicks the link, they are directed to the NotesLink webpage and review the patient’s complete medication history as well as discrepancies identified by the system, if there are any. For each patient visit, the NotesLink button is accessible via LMR for a total of 21 days, which includes 7 days prior to the patient’s scheduled clinic visit and the 14 days after.

Figure 4. The NotesLink Button in LMR’s Medication Page

NotesLink’s applies a knowledge base to identify the current status of the medication by analyzing the context surrounding it. In order for NotesLink to successfully compare the
medication lists to the free-text notes, it must assign a singular status to every medication encountered within the notes despite the numerous mentions that each medication has within a patient’s entire notes. The current logic for determining the final status of a medication mainly targets the most recent note entered within six months of the patient’s scheduled clinical visit. Three types of discrepancies are alerted by NotesLink: Possible Missing Medications, Possible Reactivated Medications, and Possible Discontinued Medications (Figure 4). “Possible Missing Medications” list presents medications that the patient may be taking, but which are not documented on the active medication list in LMR. “Possible Reactivated Medications” list presents medications that have been discontinued from the medication list in LMR, but may have been restarted. “Possible Discontinued Medications” list presents medications documented in LMR, which the patient may no longer be taking. The above sections provide decision support to care providers to manage patient’s medication list.

NotesLink also displays several important informational sections. “Considered or Requested Medications” list presents medications that either being considered or recommended by a provider or have been requested by the patient. “Active Meds Found in Notes” list presents the medications that are currently documented on the patients’ medication list in LMR. “Inactive Medications Found in Notes” list presents medications that are currently discontinued from the medication list in LMR. “Other Medications Found in Notes” list presents other medications found in notes that were never documented on the medication list in LMR and are currently not being taken by the patient. NotesLink also displays “Drug allergies found in Notes”.

Figure 5. NotesLink Key Functions

When the user clicks a medication listed in any of the above sections, NotesLink displays all patient notes that contain the selected medication. It displays a summary profile for each note, including: 1) the note subject, date, author and site; 2) the sentence containing the selected medication; and 3) the section name in which the drug was found. In addition, it also provides a hyperlink to the full note as well as advanced search and filtering functions for the notes. Clinicians can click the “reviewed” button after they reviewed the medication discrepancies to remove that medication from the list.
Implementation and Pilot

The goal of this step was to conduct a pilot and feasibility study that will stimulate future larger scale implementation or avenues for additional research in this area. We are piloting the NotesLink system in the two BWPC clinics since September 2013. In order to ensure that all clinicians undergo adequate training on how to use this tool, instructional sessions were held to demonstrate the system and provided an opportunity for hands-on practice. Reminders were sent to clinicians after the system was implemented and at regular intervals thereafter. We also created a one-page “NotesLink Quick Guide” and distributed copies of this guide to the two clinics.

SYSTEM EVALUATION AND FINDINGS

Our evaluation focused on the feasibility and performance of the proposed methods and system in improving the process of medication reconciliation in the outpatient setting. We gathered feedback on the tool from clinicians and generalized lessons learned on the usability of the tool and its impact on patient safety. To assess the system performance, we selected two randomized patients’ samples who visited the two clinics between December 1, 2013 and May 31, 2014. First, we selected 30 patients who had at least one note within the past 6 months. This set was used for a comprehensive evaluation of NotesLink, including assessing how well NotesLink identified medication information and medication discrepancies from the notes. In order to create gold standard, we used Protégé Knowtator to annotate each note for the selected patients. Two pharmacists manually reviewed all patients’ notes, annotated medication names and their status (missing, reactivated, discontinued, considered, active inactive, other, or allergy). They also identified what medication discrepancies each patient had. Any disagreements were discussed and finalized with the principle investigator. Agreement between the two pharmacist reviewers was evaluated with Cohen’s Kappa. Standard metrics including precision and recall and F-measure were computed.

In addition, 25 patients with at least one missing, discontinued, or reactivated medication discrepancy identified by NotesLink were randomly selected for a sub-analysis to identify and characterize the causes of false positive alerts. A physician and pharmacist conducted manual review to determine whether NotesLink’s categorization for each alerting medication was correct. The inter-rater agreement was calculated and the false positive rate was determined.

Our comprehensive analysis of the evaluation results is still on-going. In this report, we present some preliminary findings. Among the 30 patients, there were 1098 medication mentions (306 unique medications) found in 163 notes. Agreement between the reviewers for the annotation of medication mentions was 90.4% and for the medication status classification was 90.1% (86.0-94.2%). NotesLink achieved a precision of 70.1% and a recall of 73.7% on classifying all medication discrepancies and status (F-Measure= 71.9%). Among the 25 patients, 52 medication discrepancies were identified and 50% of these were true discrepancies according to the manual review.

For the qualitative analysis, we have conducted clinician interviews at each site. We interviewed a total of 11 clinicians, including 4 physicians, 2 pharmacist, 4 nurses and 1 physician assistant. A detailed analysis of the interviews data is still ongoing. Our preliminary
analysis showed that in general, NLP technology has not been widely used in real-time practice, though NLP tools such as NotesLink are beneficial and useful, particularly for patients with a large numbers of medications. Clinicians also suggested adding more actionable buttons to NotesLink (e.g., add or remove a medication), which can be used to edit medication list directly. Conceptually, we could implement these functions but we lack resources for integration within LMR as Partners is currently replacing it with a commercial EHR system.

**DISCUSSION**

In this study, we designed a generic medication management system architecture and developed a real-time application that combines an NLP system and a web-based user interface. The application was designed so that it can be used by different types of health care providers attempting to manage and/or reconcile medications. Depending on practice workflow and local policies, the tool can be accessed by physicians at the point of patient care. It can also be used by nurses or physician assistants who review possible discrepancies before a patient sees their physician and then communicate with the physician about these discrepancies. Similarly, the tool can be used by pharmacists who may have dedicated time to conduct medication reconciliation tasks.

During medication reconciliation, providers are often expected to review extensive medication lists within the time constraints of a patient visit. The length of these lists, as well as the time allotted for their review, is a barrier to conducting thorough medication reconciliation. However, providers are able to recognize the value of a technology that helps to remove discontinued medications or add missing medications. We were able to create a novel medication reconciliation tool based on NLP with an accuracy rate of 50% in identifying discrepancies. When a provider was alerted within LMR by NotesLink of a possible medication discrepancy, at least one discrepancy was probably a true positive in most of the cases.

Our error analysis identified several general error reasons. One source of error was due to the lack of inclusion of the medication names in our lexicon. These medication names were over-the-counter (OTC) medications and the terms used to document them in the notes were descriptive, but often non-specific (e.g., weight lifting supplements, wellness formula, DHA supplements, and artificial tears). Another main source of error involved the assignment of the wrong medication status. Other errors were due to allergies, drug classes, multiple ingredient medications, inpatient drugs, antibiotics, abbreviations and misspellings.

The accuracy of the data presented by the NotesLink is highly dependent on the quality of the notes entered into electronic health record. Due to the high incidence of copying-and-pasting of information from previous notes, NotesLink detects a high incidence of medication reactivation, which is unlikely to be accurate. Further studies are needed to determine how copying and pasting could be identified and incorporated into the logic of NLP.

NotesLink optimization in the future would include additional development of its drug lexicon and improvement of status and context identification, possibly through the use of machine learning technology. Currently, NotesLink’s logic for determining discrepancies mainly relies on the notes within 6 months of patient’s scheduled visit. More sophisticated rules need to be developed, for example, by using temporal information or specific domain knowledge. In addition, incorporating route information associated with medication entries could help
distinguish between short term inpatient medications and long term outpatient medications. Additional studies are necessary to understand the portability of NotesLink across clinics and institutions.

Conclusion

Accurate and complete medication information at the point of care is crucial for delivery of high-quality care and prevention of adverse events. However, the medication reconciliation process is very complex. In this study, we developed a novel approach and a web-application based on NLP, information retrieval, and clinical decision support to support medication reconciliation in the ambulatory setting. Our work identified gaps and challenges of using such technologies at the point of care and also provided new scientific knowledge and insights for future development.

LIST OF PUBLICATIONS AND PRODUCTS

Conference Abstracts

Manuscript in Preparation
Integration of an NLP-based Application with an Ambulatory EHR system to Support Medication Management
Identifying Copying and Pasting Medications Using a NLP based Approach.
Identifying Opiate Medication Discrepancies Using NLP and Provider Notes

REFERENCES


31. i2b2 - NLP Shared Tasks: https://www.i2b2.org/index.html.


