MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Medication Reconciliation Using Natural Language Processing and Controlled Terminologies

James J. Cimino, Tiffani J. Bright, Jianhua Li

Department of Biomedical Informatics, Columbia University, USA

Abstract

Medication reconciliation (MR) is a process that seeks to assure that the medications a patient is supposed to take are the same as what they are actually taking. We have developed a method in which medication information (consisting of both coded data and narrative text) is extracted from twelve sources from two clinical information systems and assembled into a chronological sequence of medication history, plans, and orders that correspond to periods before, during and after a hospital admission. We use natural language processing, a controlled terminology, and a medication classification system to create matrices that can be used to determine the initiation, changes and discontinuation of medications over time. We applied the process to a set of 17 patient records and successfully abstracted and summarized the medication data. This approach has implications for efforts to improve medication history-taking, order entry, and automated auditing of patient records for quality assurance.

Keywords:
drug therapy, computer-assisted; medication errors;
medication systems, hospital; natural language processing;
terminology; medical order entry systems;
quality assurance, health care

Introduction

The medications a patient takes are not always the medications the patient is *supposed* to take. Errors and drug interactions can occur when patients are treated without full knowledge of the medications that have been previously ordered, especially at transition points such as hospital admissions, transfers and discharge.[1]. For example, Beers and colleagues found that 83% of hospital admission histories failed to record the use of at least one medication the patient claimed to use or recorded at least one medication that the patient denied using; 46% had three or more such errors [2].

The process of creating an accurate list of a patient's medications, for the purposes of resolving discrepancies and supporting accurate medication ordering, is referred to as *medication reconciliation* (MR) [3]. Many processes have been developed for MR, such as manual audits and surveys. With the increasing availability of electronic patient records, such processes can be better integrated into the clinicians' workflow, but they remain largely manual tasks [4,5].

Poon and colleagues provide an excellent review of the problem of MR and describe a pilot test of a system they developed to support creation of a preadmission medication list (PAML) [6]. Their system extracts patient medication information from four different sources and presents it to the clinician, who then creates the PAML by consolidating and reconciling the information from the various sources into a single coherent list. The system assists the user by converting medication terms into, and then grouping them by, their generic names. The PAML was successfully used at the time of hospitalization to guide admission orders and at the time of discharge to assure that previous outpatient medications, where appropriate, were continued.

The PAML used data from multiple sources, each of which encoded medication data using a different controlled terminology. One of the challenges encountered by Poon and colleagues was the reconciliation of these different terminologies. The PAML did not use narrative text sources (such as discharge summaries or clinic notes) but had it done so, they would have also had to reconcile the free-text terms found in those reports.

In this paper, we describe an approach to automatically analyze medication information from a mixture of coded and narrative text sources, and use controlled terminology to support reasoning about MR. We describe the application of our method to a set of 17 patient records and discuss ways in which the results might be exploited for automated support for medication reconciliation in a variety of clinical situations.

Materials and methods

Our approach to MR involves the collection of patient medication data from multiple coded and narrative text sources, conversion of all data into coded form (using natural language processing where necessary), and then obtaining classification information for each medication. Patient medications can then be viewed over time, grouped by class, and organized into a matrix that can be used to identify points at which medications were introduced, changed or removed from the patients' ordered medications.

Sources of patient medication information

New York Presbyterian Hospital (NYPH) makes use of two major clinical information systems: the commercial product Eclipsys XA order entry system (Eclipsys, Boca Raton, FL) and the internally developed WebCIS documentation and repository system [7]. Eclipsys allows clinicians to order medications that are initiated and discontinued at various points during a hospitalization. Clinicians also enter medication information in narrative text, as part of the patient's discharge instructions. WebCIS provides narrative medication lists as parts of clinic notes, admission notes and discharge summaries, as well as coded medications from the outpatient medication order entry system and inpatient pharmacy system.[1] Together, the sources provide medication lists that correspond to twelve points in time related to hospital admissions, as shown in Table 1.

Table 1 – Data sources for patient medication information

Data Source	System	Data Type
Prior Clinic Note	WebCIS	Narrative
Prior Outpatient Medications	WebCIS	Coded
Admission Note[2]	WebCIS	Narrative
Admission Note Plan[2]	WebCIS	Narrative
Admission Orders	Eclipsys	Coded
Admission Pharmacy Orders	WebCIS	Coded
Active Orders at Discharge	Eclipsys	Coded
Discharge Pharmacy Orders	WebCIS	Coded
Discharge Instructions	Eclipsys	Narrative
Discharge Plan	WebCIS	Narrative
Clinic Note After Discharge	WebCIS	Narrative
Outpatient Medications after Discharge	WebCIS	Coded

Natural language processing

As noted in Table 1, the medication lists from clinic notes, admission notes, discharge summaries and discharge instructions are in narrative form. In order to convert them to coded form, we use a natural language processing system called Medical Language Extraction and Encoding system (MedLEE) [8] to parse the reports, and identify medication terms. Where possible, MedLEE provides Concept Unique Identifiers (CUIs) from the US National Library of Medicine's Unified Medical Language System (UMLS) [9].

Medication classification

All coded information obtained from Eclipsys (medication orders) and WebCIS (outpatient medication orders and inpatient pharmacy orders) are coded using our internally developed Medical Entities Dictionary (MED) [10]. The MED includes UMLS CUIs for many of its terms. Thus, we were able to obtain MED Codes for all coded data and, through UMLS mappings, for narrative data for which MedLEE provided UMLS CUIs.

Medication terms in the MED are organized into a hierarchy based on the American Hospital Formulary Service (AHFS) Codes, which classify drugs according to function and purpose [11]. For each medication term (including MedLEE-coded terms), we obtained the AHFS class code or codes, based on its location or locations in the MED hierarchy.

Evaluation

We chose a convenience sample of records for patients followed by one of us (JJC) in the NYPH medical clinic. For each patient, we identified the most recent hospital admission for which a discharge summary and at least one clinic note were available. We also obtained, where available, all WebCIS and Eclipsys medication data (as described above) that were recorded prior to, during, and after the hospitalization. Data were coded with AHFS codes, as described above. The data for each admission were organized into a matrix in which each row corresponded to a data source and each column corresponded to an AHFS code. The success of capturing, coding and organizing the data were measured at each step in the process.

Results

Patient data

A total of 70 patient records were reviewed and 30 hospitalizations were identified. Thirteen hospitalizations were excluded because they lacked a discharge summary and/or at least one clinic note prior to or following the hospitalization. Data from the remaining 17 hospitalizations were extracted from WebCIS and Eclipsys; narrative text was processed with MedLEE.

In general, MedLEE was successful at identifying medication terms in the narrative text. Manual review identified approximately 30 instances where MedLEE missed medication terms in the text. It also occasionally extracted nonmedication terms, such as "medication" and "po". MedLEE did not distinguish between medications that were being ordered and those that were being discontinued.

The twelve sources contributed data to a case an average of 12.7 times (range: 7-17), with a total of 1563 medications found (91.9 per case; range: 25-159). Table 2 shows the

1 Medication orders from Eclipsys are transmitted to the pharmacy system, converted into dispensing orders and transmitted to WebCIS.

2 In cases were admission notes were not available, admitting history was obtained (if present) from the discharge summary

numbers of medications obtained from each source for all 17 patients.

Table 2 – Data obtained from 17 patient records

Data Source	Meds	Records w/Data	Meds per Patient
Prior Clinic Note	157	17	9.2
Prior Outpatient Medications	211	13	16.2
Admission Note	102	14	7.3
Admission Note Plan	41	12	3.4
Admission Orders	88	8	11.0
Admission Pharmacy Orders	152	14	10.9
Active Orders at Discharge	93	8	11.6
Discharge Pharmacy Orders	171	14	12.2
Discharge Instructions	60	7	8.6
Discharge Plan	123	16	7.7
Clinic Note After Discharge	140	16	8.8
Outpatient Medications after Discharge	225	13	17.3

Coding medication data

Of the 623 terms it identified in the narrative text sections, MedLEE provided UMLS CUIs for 545. Of the remaining 78 medications, 30 were instances of five non-medication terms ("antiinflammatory", "cream", "lotion", "lozenge", and "po"). The other 48 were instances of eight medication terms ("INH", "MVI", "Os-Cal", "asa", "darvocet", "hctz", "niacin", and "toprol") for which UMLS CUIs were readily identified using the UMLS file MRCONSO.

Thus, UMLS CUIs were obtained for 593 of the MedLEE-identified terms, representing 169 unique terms and 165 unique CUIs. Eighty-five of these CUIs (representing 359 term instances) were found in the MED. The remaining 80 CUIs (representing 234 term instances) were mapped manually to the MED.

The coded data and the MedLEE-extracted data together comprised 1563 terms, of which 1533 were coded in the MED (444 unique terms) and the remaining 30 consisted of the instances of the 5 non-medication terms identified by MedLEE. Of the 1533 MED-coded terms, AHFS codes were available for 1517 (442 unique terms). The remaining 16 terms were instances of "oxygen" and "medication" that, while technically medication terms, do not have AHFS codes. Due to the multiple hierarchy of the MED, and multiple ingredients and/or uses of the medications, 270 terms (85 unique terms) mapped to two or more (maximum five) AHFS codes.

Creating medication matrices

The patient data for each case were grouped into an average of 22.8 AHFS codes (range: 10-37), and thus, given the twelve data sources, the matrices had an average of 273 cells (range: 120-444). Not every data source had data for every AHFS code; cases had data in 88 cells, on average (range: 28-164), with an average of 1.26 terms per cell (range: 1.09-1.55). An abbreviated example of a matrix is shown in Figure 1. The full matrix contains the original data from each source (row), along with the MedLEE abstraction, if appropriate, along with the 10 to 37 AHFS class columns.

Patient #9	201204: Anticoag-ulants	240400: Cardiac Drugs	240800: Hypoten-sive Agents	280000: Central Nervous System Agents	281604: Antidep-ressants
Prior Clinic Note	coumadin	verapamil	cozaar		cymbalta
Prior Outpatient Medications	Coumadin 5 mg Tab	Verapamil 180 mg Extended Release Tablet	Losartan Potassium 100 mg Tablet	Pregabalin 50mg Capsule	
Admission Note	coumadin	verapamil	cozaar		cymbalta
Admission Note Plan	coumadin				
Admission Orders	Warfarin Sodium Oral 10 MG	Verapamil SR Oral 240 MG	Losartan Oral 50 MG		
Admission Pharmacy Orders	WARFARIN TAB 5 MG 10 MILLIGRAM	VERAPAMIL SR TAB 240 MG	LOSARTAN POTAS-SIUM TAB 50 MG		
Active Orders at Discharge		Verapamil SR Oral 240 MG	Losartan Oral 50 MG		
Discharge Pharmacy Orders		VERAPAMIL SR TAB 240 MG	LOSARTAN POTAS-SIUM TAB 50 MG		DULOX-ETINE CAP 20 MG
Discharge Instructions					cymbalta
Discharge Plan					cymbalta
Clinic Note After Discharge	coumadin	verapamil			cymbalta
Outpatient Medications after Discharge	Coumadin 5 mg Tab	Verapamil 180 mg Extended Release Tablet	Losartan Potassium 100 mg Tablet	Pregabalin 50mg Capsule	

Figure 1 - Sample medication reconciliation matrix

The abbreviated view of the matrix in Figure 1 illustrates some of the information that can be obtained from these medication summaries. In particular, it shows instances where medications, such as Pregabalin (a Central Nervous System Agent) and Cymbalta (an Antidepressant), are listed in the outpatient note or ordering system prior to admission, but then do not appear in the admission note or orders. Cymbalta, at least, eventually appears in the discharge instructions and plan, but the Pregabalin does not, although it continues to be present in the outpatient medication system, raising the issue of whether it should be discontinued.

In addition to the appearance and disappearance of medications, the dosage of medications changes over the course of care. In this example, we see admitting orders for Coumadin (an Anticoagulant), Verapamil (a Cardiac Drug), and Cozaar (a Hypotensive Agent) variously increasing or decreasing, relative to the preadmission orders, only to return to their original doses after discharge. The discharge plans and discharge medication orders do not address these medications at all.

Discussion

We have successfully extracted patient medication information from a variety of sources and organized it using a standard drug classification system to support a chronological summarization by medication class. This approach is similar to that carried out by Poon and colleagues as part of their PAML application [6] but differs in two important ways: medication class and chronology.

When the medications from a particular source were grouped into an AHFS class, we tended to find approximately one drug per class. This makes clinical sense, since few patients take more than one anticoagulant or antidepressant at one time. These small groupings may help support rapid assessment of the data to detect possible problems. Furthermore, the grouping by class (as opposed to Poon and colleagues' grouping by generic ingredient) accommodates changes in medication within a class and treats them as potentially acceptable. So, for example, the fact that a patient was taking one diuretic prior to admission, a different diuretic during the admission, and returned to the previous one after discharge may reflect an intentional change (perhaps due to restrictions in the hospital formulary).

The chronological arrangement of the information in the matrices acknowledges the fact that medication lists are created and maintained at points in time and that a list that was created a year ago may not be as valid as one created a week ago. Unlike Poon and colleagues, we have not yet created a medication reconciliation application, but we believe that when patient data are displayed in such an application, the relative age of the data will be relevant to those trying to resolve differences in medication lists.

Before the present methods can be employed in an MR application, some improvements to MedLEE and the MED will be needed. Some improvements in MedLEE parsing will be needed, particularly to allow it to differentiate between mentions of medications that are being continued and those that are being discontinued. MedLEE did extremely well at providing UMLS CUIs for the terms it did find - it missed only 8 unique terms out of 172 legitimate unique medication terms, for a 95% success rate. Our experience suggests that only a small effort will be needed to improve this rate.

The automated translation of UMLS CUIs to MED Codes was less successful, at slightly over 50%. This result is not due to the methodology but rather to the heretofore-low incentive for having UMLS CUIs in the MED; the mappings have not been updated for some time. However, we are confident (based on the success of our manual mapping) that the MED can be easily brought up to date and will perform this mapping function more successfully in the future.

Finally, while we believe that the matrices provide a useful organization of the data, they are probably not adequate for use directly by clinicians. Some applications, such as the PAML, will be needed to present the information to users in a way that reduces the cognitive overload that might otherwise be produced by a 164-cell matrix.

Alternatively, a program that attempts to identify discontinuities in patient medications, in order to alert clinicians at appropriate times, might use the information in the matrices. For example, when a patient is admitted and a previous outpatient medication is not ordered, the clinician might be asked, "The patient was previously taking X; do you wish to continue it?" The clinician might have a perfectly good reason for not continuing the drug, but inadvertently overlooking such information is unfortunately frequent [1]. Later, when the patient is discharged, the commonly used shorthand "continue all previous medications" is ambiguous - was X deliberately stopped or not? And, if deliberately stopped, is reinstatement really desired? A system that can track these discrepancies has great potential to reduce medication errors.

The approach described in this paper cannot, by itself, determine if a medication discrepancy represents a true error in medication ordering. Only discussion with the clinicians caring for the patient and/or chart review can identify the reasons for medication changes. However, if an ordering clinician overlooks a prior medication, or adds a previously undocumented one, it will be identified in our matrix because, at least for inpatient medications, the Eclipsys data are the gold-standard for the patient's intended therapy. Such identification is a necessary first step in preventing or correcting errors in medication reconciliation.

Conclusions

We have successfully extracted patient medication information from multiple systems and applications, and classified the information based on drug class. The resulting matrices of medications organized chronologically by class provide a new way to summarize such information and have the potential to support automated medication reconciliation.

Acknowledgments

The authors thank Dr. Carol Friedman for the use of MedLEE. This work was supported in part by National Library of Medicine Training Grant LM07079-11 and by National Library of Medicine Grant 5R01LM007593-05.

References

[1] Gleason, KM, Groszek JM, Sullivan C. Reconciliation of Discrepancies in Medication Histories and Admission Orders of Newly Hospitalized Patients. *Am J Health-Syst Pharm*. 2004;61:1689-95.

[2] Beers MH, Munekata M, and Storrie M. The accuracy of medication histories in the hospital medical records of elderly persons. *J Am Geriatr Soc.* 1990 Nov;38(11):1183-7.

[3] Institute of Healthcare Improvement. Medication Reconciliation Review. http://www.ihi.org/IHI/Topics/PatientSafety/MedicationSystems/Tools/Medication+Reconciliation+Review.htm
[4] Pronovost P, Weast B, Schwarz M, Wyskiel RM, Prow D, Milanovich SN, Berenholtz S, Dorman T, and Lipsett P. Medication reconciliation: a practical tool to reduce the risk of medication errors. *J Crit Care*. 2003 Dec;18(4):201-5.
[5] Poole DL, Chainakul JN, Pearson M et al. JHQ 177 Medication reconciliation: A necessity in promoting a safe hospital discharge. *NAHQ*. 2006 May-June 2006 www.nahq.org/journal/ce/article.html?article_id=256
[6] Poon EG, Blumenfeld B, Hamann C, Turchin A, Graydon-Baker E, McCarthy PC, Poikonen J, Mar P, Schnipper JL, Hallisey RK, Smith S, McCormack C, Paterno M, Coley CM, Karson A, Chueh HC, Van Putten C, Millar SG, Clapp M, Bhan I, Meyer GS, Gandhi TK, and Broverman CA. Design and implementation of an application and associated services to support interdisciplinary medication reconciliation efforts at an integrated healthcare delivery network. *J Am Med Inform Assoc*. 2006 Nov-Dec; 13(6):581-92.
[7] Hripcsak G, Cimino JJ, and Sengupta S. WebCIS: large scale deployment of a Web-based clinical information system. *Proc AMIA Symp*. 1999;:804-8.
[8] Friedman C, Shagina L, Lussier Y, Hripcsak G. Automated encoding of clinical documents based on natural language processing. *J Am Med Inform Assoc*. 2004 Sep-Oct;11(5):392-402.
[9] Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Methods Inf Med*. 1993 Aug;32(4):281-91.
[10] Cimino JJ. From data to knowledge through concept-oriented terminologies: experience with the Medical Entities Dictionary. *J Am Med Inform Assoc*. 2000 May-Jun;7(3):288-97.
[11] American Society of Health-System Pharmacists. *AHFS Drug Information*. Bethesda, MD: American Society of Health-System Pharmacists, 2006.

Address for correspondence:

James J. Cimino, M.D.
Department of Biomedical Informatics
Columbia University College of Physicians and Surgeons
622 West 168th Street, Vanderbilt Clinic 5th Floor
New York, New York 10032 USA
jjc7@columbia.edu