

Identifying Indications for Drugs in Drug Package Inserts

Chiang S. Jao, Kin Wah Fung, and Dina Demner-Fushman
National Library of Medicine, Bethesda, MD, USA

Abstract

Linking drugs and their indications can help improve the quality of electronic medical records (EMRs) and facilitate clinical decision support. We investigate an automatic approach to extraction of indications for marketed drugs from drug package inserts. Our aim is to build a reference table that associates medications and their indications. This resource could facilitate maintenance of medication and problem lists in the EMR. We are developing modules for downloading drug package inserts from the DailyMed site, parsing the XML file for each drug, and submitting the resulting texts to MetaMap. All relevant information will be represented by codes from standard terminology systems.

Introduction

Accurate and well-managed medication and problem lists are critical to success of Clinical Decision Support systems, pharmacovigilance, and other patient safety efforts. [1] Medication reconciliation depends on an accurate and up-to-date medication list. Previous research has shown that problem list and medication list data can be used to mutually validate each other. [2] One prerequisite to using this approach is to know what drugs are prescribed for what problems. Previous indications lists have been mostly hand-crafted, which are tedious to create and maintain. We are exploring a novel way to automatically extract drug indications from authoritative information in the FDA approved package inserts. Our aim is to establish a reference table linking drugs and their indications using identifiers from standard terminologies.

Methods

The DailyMed website provides access to thousands of package inserts. Each drug is linked to its RxNorm identifier (RXCUI). [3] We are developing modules for: (1) downloading drug package insert information in the Extensible Markup Language (XML) format available from the DailyMed website; (2) parsing these XML files and extracting free text in the indications section which is identified by the LOINC code 34067-9 FDA package insert Indications & usage section (3) dividing the indication text into smaller chunks making use of XML tags (e.g. titles, paragraphs or list items) to facilitate subsequent processing; (4) processing and identifying the indications concepts (represented by their Unified

Medical Language System (UMLS®) concept unique identifiers (CUIs)) using MetaMap [4].

Results

We tested a pool of 329 of the most frequently prescribed mono-ingredients drugs. The average size of indication sections in the sample files is 1176 characters. Further subdivision of these yielded 2242 text chunks. The chunking step allows to better situate the extracted concepts in the context of the insert. This information will be used in selecting final intervention candidates for a manual review of the reference table. The average size of these text chunks is 210 characters. The observed MetaMap processing time is significantly related to the number of indication terms within the chunk, rather than the size of the chunk file. In less than 1% of testing samples there were no UMLS concepts identified.

Discussion

We started exploration of automatic methods for building a reference table containing indications for marketed drugs. The relations extracted from the drug package inserts are stored as pairs of drug RXCUIs [5] and CUIs of the associated indications. Our next step is to validate the results by comparing them with other publicly available sources of similar information e.g. NDF-RT (National Drug File Reference Terms) [6], and refine the algorithm accordingly.

References

1. Joint Commission. 2010 National Patient Safety Goals for Hospitals. Available: <http://www.jointcommission.org/PatientSafety/NationalPatientSafetyGoals/http://www.nlm.nih.gov/research/umls/rxnorm/>.
2. Jao CS, Hier DB, Galanter WL. Using clinical decision support to maintain medication and problem lists: A pilot study to yield higher patient safety. 2008 IEEE International Conference on Systems, Man and Cybernetics, 10-16 October 2008, Singapore, 739-743.
3. <http://dailymed.nlm.nih.gov/>.
4. <http://mmtx.nlm.nih.gov/>.
5. <http://www.nlm.nih.gov/research/umls/rxnorm/overview.html>.
6. <http://www.nlm.nih.gov/research/umls/sourcerel easedocs/2008AB/NDFRT/>.