A Knowledge Intensive Approach to Mapping Clinical Narrative to LOINC

Marcelo Fiszman, M.D., Ph.D., Dongwook Shin, Ph.D., Charles A. Sneiderman, M.D., Ph.D., Honglan Jin, Ph.D., Thomas C. Rindflesch, Ph.D.
Lister Hill National Center for Biomedical Communications, National Library of Medicine, Bethesda MD 20894

Abstract

Many natural language processing systems are being applied to clinical text, yet practical results are obtained only by honing a system to a particular context. We suggest that concentration on the information needed for this processing is crucial and present a knowledge intensive methodology for mapping clinical text to LOINC. The system takes published case reports as input and maps vital signs and body measurements and reports of diagnostic procedures to fully specified LOINC codes. Three kinds of knowledge are exploited: textual, ontological, and pragmatic (including information about physiology and the clinical process). Evaluation on 4809 sentences yielded precision of 89% and recall of 93% (F-score 0.91). Our method could form the basis for a system to provide semi-automated help to human coders.

Introduction

Automatic processing of clinical text needs to be of high quality to be useful. There are numerous systems based on a variety of approaches addressing an array of clinical applications [1]. The most mature, general system is MedLEE [2]. So far, effectiveness in natural language processing for clinical text is of acceptable quality only by crafting a system for a particular context (e.g. [3] and the systems referenced). The main challenge is to resolve ambiguity, which requires knowledge, and this problem has not been faced in a principled way. Systems, both rule-based and statistical, look for patterns in text, but rarely categorize the type of information needed to resolve ambiguity [3].

We are devising a methodology for mapping clinical text to LOINC (Logical Observation Identifier Names) codes [4] and present preliminary results from a knowledge intensive system that takes published case reports as input and maps vital signs and body measurements and diagnostic procedures to fully specified LOINC codes. We have begun to overtly categorize the type of knowledge needed into linguistic, ontological (from UMLS), and pragmatic (e.g. about document structure, physiology, and clinical process).

LOINC provides a suitable challenge for testing a knowledge intensive approach. The extensive detail represented in a LOINC code forces a system to provide appropriate information. A practical consequence of an effective system would be to provide semi-automated assistance for human coders.

Background

LOINC

LOINC is a structured vocabulary for encoding laboratory tests and clinical observations [5]. "If we consider the observation as a question and the observation values as answers, LOINC provides codes for the questions." [6] Previous research has not attempted to map clinical text to LOINC, but rather has focused on mapping local terminologies; for example Intelligent Mapper [7] maps local radiology terms to LOINC. Other lexical approaches include [8,9]. Additional work investigates semantic interoperability between LOINC and other standard terminologies (e.g. [10,11]).

Concepts in LOINC have a mnemonic "long common name" and a fully specified code as a synonym. Each fully specified code has five or six fields that serve as building blocks for recording detailed descriptions of laboratory tests and clinical observations, including diagnostic procedures . For example, "Blood pressure systolic & diastolic" is the long common name for the fully specified code in (1), which provides a framework for reporting a clinical observation and comprises Field 1, component ("Blood pressure systolic & diastolic"); Field 2, property ("Pressure"); Field 3, timing ("Point in time"); Field 4, system ("Arterial system"); and Field 5, type of scale ("Quantitative").

1. Blood pressure systolic & diastolic:Pressure:Point in time:Arterial system:Quantitative

Additionally, the method used, although optional, is reported in Field 6 when appropriate, as for radiology diagnostic procedures. For example, in (2b) (with long common name (2a)), the value of Field 6 is "Computerized Tomography." The value of Field 1 ("Multisection") represents projections and spatial

conditions when Field 6 is a radiology diagnostic procedure.

2. a. Head CT
   b. Multisection:Finding:Point in time:Head:Narrative:Computerized Tomography

LOINC in the UMLS Metathesaurus

The LOINC vocabulary is included in the Unified Medical Language System® (UMLS)® Metathesaurus [12]. The individual fields and the fully specified code are represented separately; the relationship between a field and the fully specified code it is a part of is represented with the Metathesaurus concept attribute "Other Relations." The relationship between the individual field and the fully specified code reflects LOINC implied semantic relations. For example, the concept "Computerized Tomography" (from the LOINC code above) is represented in the UMLS and has relations to all fully specified LOINC codes in which "Computerized Tomography" appears in Field 6. This relationship is encoded with the Metathesaurus attribute "Other Relations" having value "method_of" for the concept "Computerized Tomography." We exploit this Metathesaurus characteristic when using MetaMap and the UMLS Knowledge Source Server (UMLSKS) [13] to find the best LOINC code to represent phenomena in published case reports.

MetaMap

The MetaMap application [14] maps biomedical text to concepts in the UMLS Metathesaurus. During processing, MetaMap provides an underspecified syntactic analysis, which breaks input into phrases. For example, the structure represented schematically in (4) is produced for input (3).

3. Admission vital signs included … blood pressure of 92/53 mm Hg, …

4. [Admission vital signs] [included] … [blood pressure] [of 92/53 mm Hg,] …

For each phrase in a structure such as (4), MetaMap returns the best Metathesaurus concepts, including multiple concepts due to ambiguity. Semantic types are returned with all concepts. MetaMap returns the mappings in (5) for *blood pressure*.

5. Blood Pressure [Organism Function]
   Blood pressure finding [Finding]
   Systemic arterial pressure [Finding]

MetaMap can be configured so that it only returns concepts in a specified UMLS component vocabulary, such as LOINC.

Methods

Overview

Case reports were taken from The New England Journal of Medicine, due to the clinical diversity presented in this journal. Example titles include, "A 20-year-old woman with asthma and cardiorespiratory arrest" (PMID 17507708), "A 61-year-old man with a mediastinal mass" (PMID 17522403), and "A 19-year-old college student with fever and joint pain" (PMID 17582074). Typical text in the reports is: *Admission vital signs included a fever of 103.3°F, blood pressure of 92/53 mm Hg (systolic/diastolic), pulse of 84 beats per minute, respiration of 22 breaths per minute, and oxygen saturation of 95% on bilevel positive airway pressure (BiPAP) ventilatory support. A portable bedside chest radiograph obtained at the initial evaluation showed peripheral patchy opacities.* We randomly, retrieved 140 reports and automatically extracted the case presentations (approximately 65 sentences per presentation). Sixty-five reports (4105 sentences) were used for training and 75 (4809 sentences) for testing.

We selected phenomena commonly reported in clinical text to be mapped to LOINC. These are in two categories: vital signs and body measurements and common diagnostic procedures. For radiology diagnostic procedures we also identify the target body location. There is a separate LOINC code for each radiology diagnostic procedure and the body location (system) to which it pertains. Long common names for the 23 LOINC codes we retrieve are given in (6a) and (6b) We currently do not retrieve measurement values nor the results of diagnostic procedures (i.e. we capture LOINC "questions").

6. a. Blood pressure systolic & diastolic, Heart rate, Respiratory rate, Body temperature, Body weight, Body height Oxygen saturation in Blood
   b. Cardiac echocardiogram study, Abdomen CT, Chest CT, Head CT, Pelvis CT, EKG, Abdomen MRI, Chest MRI, Pelvis MRI, Brain MRI, Heart MRI, Abdomen US, Chest US, Pelvis US, Abdomen X-ray, Chest X-ray

Our method for mapping concepts in clinical text to LOINC comprises three parts: textual analysis, which includes preprocessing and calling MetaMap; accessing the LOINC vocabulary with UMLSKS; and LOINC-specific processing, which identifies the best LOINC code for the input text. The algorithm is essentially an extension of the "lexically assign,

logically refine" principle [15], in that additional knowledge enhances initial string matching.

Knowledge

Three kinds of knowledge are exploited: textual, ontological, and pragmatic. Linguistic knowledge includes information derived from textual context, while the Metathesaurus provides ontological knowledge, especially in the form of semantic types. Pragmatic knowledge is of various kinds and derived from various sources (currently made available in the form of rules). Medical knowledge includes characteristics of health and disease (Physiologic functions are processes carried out by organs, tissues, and cells to maintain health.) Knowledge about the clinical process includes information about document type and structure as well as extensive knowledge held by health care professions, such as the default modalities for diagnostic procedures and measuring vital signs and body attributes, which are assumed if nothing else is given.

Mapping to appropriate LOINC codes requires knowledge about the kind of information encoded and how it is represented. For example, it is necessary to know that vital signs are reported with quantitative values and that some observations refer to the patient generally, rather than to a particular body system.

Preprocess text

Before MetaMap is called, instances of *pulse* in the text are replaced with *heart rate,* if *pulse* is followed by an integer with no more than two words intervening (e.g. *pulse of 84 beats per minute*, but not *pulse was 2+*). This is necessary because these terms are not synonyms in the Metathesaurus. Similarly, we substitute *respiratory rate* for *respiration* (also when followed by an integer with no more than two words intervening, e.g. *respiration of 22 breaths per minute*). These substitutions are based on the knowledge that this usage is common in clinical text, even though, strictly, neither respiratory rate and respiration nor pulse and heart rate are equivalent.

Based on the knowledge that the textual context is presentation of the examination facts of a clinical case, we expand acronyms *CT* to *computed tomography* and *US* to *ultrasound.* There are many proposals for expanding abbreviations in clinical text (e.g. [16]), which we have so far not exploited.

MetaMap

After preprocessing, MetaMap limited to LOINC is run on input text noun phrases to provide concepts and associated semantic types. The ontological knowledge inherent in semantic types guides further processing. Vital signs and body measurements have one of the following semantic types: 'Clinical Attribute', 'Organism Function', 'Organism Attribute', or 'Organism or Tissue Function'. The semantic type 'Diagnostic Procedure' refers to diagnostic procedures. LOINC codes for radiology diagnostic procedures require an associated body location, which either precedes or follows the procedure name in text and can be coordinated. Noun phrases mapping to concepts with semantic type 'Body Location or Region', 'Body Part, Organ, or Organ Component', or 'Body Space or Junction' refer to body locations.

This output from MetaMap is the LOINC code part which best matches input lexically. For example, text *computed tomography* matches LOINC code part "CT." These concepts are ontologically consistent with the code appropriate for the input text, but are generic and do not provide necessary detail.

Knowledge Source Server

UMLSKS is called to get all fully specified LOINC codes which have the generic LOINC code part as the value of some field. Knowledge about the way LOINC is represented in the Metathesaurus is exploited during this processing. The Metathesaurus concept attribute "Other Relations" is used for this purpose. Vital signs and body measurements have "has_component" as the value of this attribute. Diagnostic procedures have "method_of." For example, the generic LOINC code part "CT" is related to 696 fully specified LOINC codes through this attribute. The appropriate final code is selected from these through a winnowing process that takes advantage of knowledge about the information LOINC encodes and how it is represented.

LOINC processing

In applying knowledge-based rules to select the best code, we initially assume that observations reported in a case report reflect aspects of the default clinical process and eliminate codes which refer to special circumstances. With the exception of the string "^Patient," we eliminate any code that has period (".") or up-arrow ("^") in any field. The period normally indicates a more specific code and the up-arrow usually indicates some particular circumstance in which a physiologic function was measured, such as "^pre transfusion" or "^post phlebotomy." For example, (7) is eliminated on this basis, based on the knowledge that if it isn't specified, the blood pressure measurement was not done after a transfusion.

7. Blood pressure systolic & diastolic^post transfusion:Pressure:Point in time:Arterial system:Quantitative

Different rules are applied to vital signs and body measurements on the one hand and diagnostic procedures on the other. For vital signs and body measurements, in general, processing prefers codes with exact value "^Patient" in Field 4 (system) of the fully specified LOINC code. This rule is based on the knowledge that most such measurements refer to the patient generally. Further, codes for vital signs and body measurements have "Quantitative" as the value of the Field 5 (scale).

For blood pressure, several rules apply. Field 4 (system) has value "Arterial system" based on knowledge of physiology that blood pressure is an arterial cardiovascular physiologic function. Field 2 (property) has value "Pressure" and not "Pressure Difference," based on the pragmatic clinical knowledge that if not specified, a blood pressure measurement indicates a single value, not a comparison of measurement values over time. In this case, the structure of LOINC requires that Field 3 (temporal aspect) have value "Point in time." After applying these rules, (8) remains as the single best code for *blood pressure* when it occurs in published clinical case reports.

8. Blood pressure systolic & diastolic:Pressure:Point in time:Arterial system:Quantitative

All diagnostic procedure codes have "Narrative" as the value of 5th field (scale). Based on procedure characteristics, Field 1 (component) in LOINC codes for radiology diagnostic procedures represents projections and spatial conditions. This needs to be stipulated for each procedure. For example, for computed tomography and MRI this value is "Multisection," while for radiograph it is "Views." Diagnostic procedures are executed on various parts of the body, and the value of Field 4 (system) of diagnostic procedure codes represents the pertinent location. The system identifies this specification in input text and associates it with mention of the procedure. Applying these rules, two fully specified codes (9) are retrieved for input text *Computed tomography of the chest and abdomen*.

9. a. Multisection:Finding:Point in time:Chest:Narrative:CT
b. Multisection:Finding:Point in time:Abdomen:Narrative:CT

Evaluation

One of the authors (CAS), a board certified family medicine physician not involved in system development, annotated the 75 reports selected for testing with the 23 selected LOINC codes. We then compared system output to this reference standard and calculated recall, precision, and F-score.

Results

The reference standard contained 878 annotated codes, and the system produced 917 codes, 813 of which were true positives, for precision of 89%. There were 65 false negatives, for recall of 93%. The F-score was 0.91. We also examined results for vital signs and body measurements separately from diagnostic procedures. There were 500 codes annotated vital signs and body measurements, and the system produced 535 codes, 475 of which were true positives, for precision of 89%. There were 25 false negative, for recall of 95% (F-score 0.91). There were 378 codes marked as diagnostic procedures, and the system produced 382, of which 338 were true positives, for precision of 88%. Forty were false negatives, for recall of 89% (F-score 0.89).

Discussion

Error analysis indicated that false positives were mostly due to a comment on a characteristic of a vital sign, rather than reporting quantitative value, as in (10).

10. A 10-year-old girl was seen…because of dyspnea and noisy respiratory rate

This could be corrected by requiring that vital signs must be followed by a number if they are to be mapped to a quantitative LOINC code.

The majority of false negatives were due to anaphora, which we currently do not address. For example, the text (12) closely followed (11) in a report.

11. Ten months before this evaluation, CT of the abdomen and pelvis revealed …

12. Three months before this evaluation, repeat CT revealed…

In the reference standard, both (11) and (12) were annotated with the codes for CT of the abdomen and CT of the pelvis, based on the fact of (12) succeeding (11) in the report. The appropriate codes for (11) were retrieved by the system, but not those for (12).

The results based on knowledge intensive processing appear to be promising for mapping published case reports to some classes of LOINC codes. Although this text is not identical to clinical narrative, it likely approximates some parts of the patient record, especially discharge summaries.

We annotated only a subset of codes for radiology diagnostic procedures, namely those for commonly targeted body locations. This limitation was largely in consideration of annotator burden. The text

processing identifies all body locations that are the target of an imaging procedure. Additional extensions appear to be feasible. The principles exploited for vital signs, body measurements, and diagnostic procedures should also apply at least to the common laboratory tests. The method appears to have considerable promise in providing semi-automated assistance to human LOINC coders, who would make final decisions based on a limited number of options presented by the system, for example for choosing Field 1for radiology diagnostic procedures.

Conclusion

We present a knowledge intensive system for mapping vital signs, body measurements, and reports of diagnostic procedures found in published case reports to fully specified LOINC codes. The system comprises three parts: textual analysis, which includes preprocessing and calling MetaMap; accessing the LOINC vocabulary with the UMLS Knowledge Source Server; and LOINC-specific processing, which identifies the best LOINC code for the input text. Three kinds of knowledge are exploited: textual, ontological (UMLS semantic type), and pragmatic (including information about physiology and the clinical process). The case presentations from 75 reports from The New England Journal of Medicine were evaluated, yielding 89% precision and 93% recall (F-score 0.91). We suggest that the system could form the basis for providing semi-automated help to human coders.

Acknowledgements

References

1. Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF. Extracting information from textual documents in the electronic health record: a review of recent research. Yearb Med Inform. 2008:128-44.
2. Friedman C. A broad-coverage natural language processing system.Proc AMIA Symp. 2000:270-4.
3. Uzuner O. Recognizing obesity and comorbidities in sparse data. J Am Med Inform Assoc. 2009 Jul-Aug;16(4):561-70.
4. Forrey AW, McDonald CJ, DeMoor G, Huff SM, Leavelle D, Leland D, Fiers T, Charles L, Griffin B, Stalling F, Tullis A, Hutchins K, Baenziger J. Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. Clin Chem. 1996 Jan;42(1):81-90.
5. Huff SM, Rocha RA, McDonald CJ, De Moor GJ, Fiers T, Bidgood WD Jr, Forrey AW, Francis WG, Tracy WR, Leavelle D, Stalling F, Griffin B, Maloney P, Leland D, Charles L, Hutchins K, Baenziger J. Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary. J Am Med Inform Assoc. 1998 May-Jun;5(3):276-92.
6. McDonald CJ, Huff SM, Suico JG, Hill G, Leavelle D, Aller R, Forrey A, Mercer K, DeMoor G, Hook J, Williams W, Case J, Maloney P. LOINC, a universal standard for identifying laboratory observations: a 5-year update. Clin Chem. 2003 Apr;49(4):624-33.
7. Vreeman DJ, McDonald CJ. A comparison of Intelligent Mapper and document similarity scores for mapping local radiology terms to LOINC. AMIA Annu Symp Proc. 2006:809-13.
8. Sun JY, Sun Y. A system for automated lexical mapping. J Am Med Inform Assoc. 2006 May-Jun;13(3):334-43.
9. Khan AN, Griffith SP, Moore C, Russell D, Rosario AC Jr, Bertolli J. Standardizing laboratory data by mapping to LOINC. J Am Med Inform Assoc. 2006 May-Jun;13(3):353-5.
10. Bodenreider O. Issues in mapping LOINC laboratory tests to SNOMED CT. AMIA Annu Symp Proc. 2008 Nov 6:51-5.
11. Rossi Mori A, Consorti F. Exploiting the terminological approach from CEN/TC251 and GALEN to support semantic interoperability of healthcare record systems. Int J Med Inform. 1998 Feb;48(1-3):111-24.
12. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res. 2004 Jan 1;32(Database issue):D267-70.
13. Thorn KE, Bangalore AK, Browne AC. The UMLS Knowledge Source Server: an experience in Web 2.0 technologies. AMIA Annu Symp Proc. 2007 Oct 11:721-5.
14. Aronson AR. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. Proc AMIA Symp. 2001;17-21.
15. Dolin RH, Huff SM, Rocha RA, Spackman KA, Campbell KE. Evaluation of a "lexically assign, logically refine" strategy for semi-automated integration of overlapping terminologies. J Am Med Inform Assoc. 1998 Mar- Apr;5(2):203- 13.
16. Xu H, Stetson PD, Friedman C. A study of abbreviations in clinical notes. AMIA Annu Symp Proc. 2007 Oct 11:821-5.