Clinical Application of the UMLS in a Computerized Order Entry and Decision-Support System

Antoine Geissbuhler, MD., Randolph A. Miller, MD.
Division of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN

ABSTRACT

Vanderbilt University Medical Center uses the UMLS as a dictionary, an interlingua, and a knowledge source within the WizOrder system. WizOrder provides direct care-provider order entry and integrated clinical decision-support capabilities. Linking the two functions enables efficient decision-support during the "normal" workflow of care providers, at the point where decisions are made. WizOrder uses the UMLS as a dictionary to encode free-text entries into controlled vocabularies such as ICD9. As an interlingua, the UMLS provides mapping between vocabularies, allowing to translate patient-specific information to MeSH terms and perform automated literature retrieval. WizOrder uses the tables of co-occurring concepts and the Semantic Network to provide sensible lists of potential drug interactions and adverse drug reactions, and generate fully-formed MEDLINE queries for PubMed.

INTRODUCTION

This paper describes application of the Unified Medical Language System (UMLS) in daily clinical practice at Vanderbilt University Medical Center (VUMC). At VUMC, UMLS resources serve three functions: as a dictionary, interlingua, and knowledge source. These functions support the WizOrder clinician order entry and decision support system1,2, developed and implemented at VUMC. WizOrder is routinely used by physicians, nurses, medical receptionists, and other care providers for entry of all inpatient orders at Vanderbilt University Hospital (excluding two pediatric intensive care units). Approximately 6500 orders are entered and processed daily through WizOrder, which services a total of approximately 550 implemented hospital beds.

Writing medical orders represents a critical moment in the clinical decision-making process: it is the point where decisions are finalized and turned into actions through communication with care providers, clinical departments, and ancillary services. The ability to intervene at this time through providing computerized decision-support capabilities has been shown to be beneficial both for the quality3 and efficiency4 of patient care.

An essential aspect of computerized order entry involves the translation of the free-text orders formulated by care providers into structured data elements using a controlled vocabulary, for storing information in databases and communicating with other automated systems such as scheduling, billing, and decision-support systems. The UMLS Metathesaurus5 facilitates mapping free-text entries to controlled vocabularies6. As the Metathesaurus contains a variety of biomedical terminology systems, and provides links between them, it can also be used as an interlingua, allowing the integration of functions that use different nomenclatures. The UMLS Knowledge Sources, including the Semantic Network and databases such as the co-occurrence of concepts in the MEDLINE database, provide meaningful clinical knowledge related to order entry and decision support.

UMLS AS A DICTIONARY WITHIN WIZORDER

A key component of providing timely decision support is the ability to recognize the context in which clinical decisions are made. For example, during routine admission order generation, physicians traditionally enter orders stating the patient's diagnosis, condition, and diet (among many other items). Ideally, these data should be captured in structured vocabularies rather than free text. Rosenberg7 showed that the UMLS can be used effectively for coding unconstrained physician diagnoses. Various investigators have developed
applications that map free-text entries to UMLS concepts, using either lexical manipulations\(^8\text{-}^{10}\) or natural language processing\(^11\).

The ICD9 nomenclature has long been used at our institution to encode diagnoses and surgical procedures. We use the completer of the WizOrder system, which supports synonyms, abbreviations and some degree of spelling error correction, to map free-text entries of diagnoses and procedures to the lexical variants of the ICD9 keywords defined in the UMLS Metathesaurus (Fig. 1). Within the "diagnosis" order, the completer displays a ranked list of potential ICD9 codes matching the user's free-text entry, and enables the care provider to select the proper one. If the number of potential matches is excessive, the list is limited to higher-level codes, defined by the numeric representation of ICD9 codes. When such a "parent" code is selected, it is exploded into its "children", so that the user can select a more granular code, whenever appropriate. It is important to note that this encoding task is integrated in the normal order entry process and therefore only represents a minimal distraction. Most medications (94%) described in the institution's dictionary of orderable items are automatically mapped to UMLS Metathesaurus concepts (Concepts Unique Indentifiers, or CUIs) using their generic name. The other medications (i.e., new substances not represented in the Metathesaurus, or combinations of substances), as well as a subset of the radiological procedures and laboratory tests, are coded via pre-stored local manual mappings. These codes are used to facilitate decision-support capabilities described below. Note that the mapping to generic substance names does not preserve information about compound or proprietary medication formulations, or routes of administration: for example, Ecotrin\(^\text{®}\), an enteric-coated form of aspirin, is represented in the UMLS as a lexical variant of aspirin.

**UMLS AS AN INTERLINGUA WITHIN WIZORDER**

Since the UMLS Metathesaurus links terms from several nomenclatures (Fig. 1), it facilitates the conversion of terms from one controlled medical vocabulary to another\(^6\text{-}^{12}\). Previous investigators have developed a number of programs to map free-text to concepts and concepts to Medical Subject Headings (MeSH), in order to build MEDLINE queries\(^10,13,14\). One of the challenges of retrieving patient-specific bibliographic references is to define the context of the query. As noted above, it is the custom of physicians in writing admission and subsequent orders to describe components of the patient's clinical context. In WizOrder, the list of diagnoses, procedures, prescribed medications, laboratory tests

![Diagram](image-url)

**Figure 1.** Some of the lexical variants (terms) and strings that are grouped as the concept "Influenza", as well as links to various nomenclatures, in the UMLS Metathesaurus.
and radiology exams are readily available. They are coded from "free-text" (local lexicon) entries into CUIs or ICD9 codes, which in turn correspond to CUIs, and can be mapped to MeSH terms straightforwardly using the UMLS Metathesaurus, assuming that the given concept has both links to ICD9 and to MeSH.

It is logical to use the orders entered into WizOrder to extract a set of MeSH terms to generate a MEDLINE query by combining all the terms using a logical "OR" operator. Using all codes derived from the patient's orders indiscriminately is equivalent to the free-text query "what might I read in the indexed literature about this patient?". When using the same list of codes plus (ANDeD with) an arbitrary CUI, e.g., a new medication, the query becomes "what is potentially relevant about this medication for this patient?". Currently, for efficiency, WizOrder does not run a query against the whole MEDLINE database, but against a subset of about 30,000 citations from the last five years selected from 25 major medical journals identified by Vanderbilt clinicians as being most frequently relevant to patient care. The subset is updated monthly using the VUMC library MEDLINE query software. MeSH descriptors, citations, and abstracts for those journals are stored locally within WizOrder servers and indexed via MeSH. References, retrieved locally within a few seconds, are ranked according to the quality of the match (matching more terms from the chart simultaneously). After each query, users can selectively display and/or print citations, including abstracts, after reviewing a summary list of authors, titles, and journal sources. This simple search strategy only matches identical MeSH terms, and does not use the MeSH hierarchy. This can be problematic when MeSH terms derived from the patient's record are very specific. Therefore, the scope of the query is automatically broadened when too few references are retrieved: drugs are replaced by their therapeutic class (e.g., furosemide is replaced by loop-diuretic), and diagnoses by their "parent" code in the ICD9 hierarchy (e.g., atrial flutter, code 427.32, is replaced by cardiac dysrhythmia, code 427). Users can further focus or refine queries by manually adding arbitrary MeSH terms to the list of those that were automatically extracted. These additional terms are "ANDeD" with existing terms.

Instead of combining all MeSH terms generated from the chart, more focused queries based on pairs of MeSH terms can be generated. Further refinement of queries using the UMLS Semantic Net, the MeSH subheadings and the MEDLINE concepts co-occurrence information is described in the following section. Canned MeSH queries, such as the ones developed in the QMR knowledge base\textsuperscript{15}, will also be used to improve the performance of the function.

**UMLS AS A KNOWLEDGE BASE WITHIN WIZORDER**

The co-occurrence of concepts in the biomedical literature represents a form of knowledge base, under the assumption that if two concepts are used in the same journal article, they must somehow be related. The number of citations for a given co-occurring pair of concepts can be an indication of the strength of their link. Applications have used these assumptions to evaluate the relevancy of concepts, using the UMLS table of co-occurring concepts in MEDLINE citations\textsuperscript{9,10}.

The WizOrder system uses co-occurring concepts as a method to generate plausible lists of drug-drug and drug-diagnosis interactions. This feature is a complement to the more deterministic drug-drug interaction database manually maintained by the VUMC pharmacy and made available in WizOrder. Each pair of co-occurring concepts in the UMLS tables is examined (Fig. 2). The Semantic Network is used to reject concepts that are neither a "Pharmacologic Substance" nor a "Disease or Syndrome". The Related Concepts table is used to determine if the relation between the two concepts is valid, e.g., that the "Pharmacologic Substance" causes, or has as a complication the "Disease or Syndrome". As the frequency of co-occurring concepts in the UMLS Metathesaurus is specified at the concept's subheading level, it is possible to estimate the quality of the co-occurrence by applying heuristics that measure how the frequency of specific subheadings such as AE (adverse effects), CT (contra-indications) and TO (toxicity) relate to the overall frequency of the co-occurrence of the concepts. This process yields a table of likely significant drug-drug and drug-diagnosis interactions. As described above, both diagnoses and medications are encoded as CUIs during the order-entry process. It is therefore straightforward to lookup the co-occurrence information and issue warning messages, that, although non-deterministic, can indicate that "there are several citations in MEDLINE of (drug X) having adverse effects described along with (drug/diagnosis Y)". The wording of the statement does not imply that the two concepts are causally related, only that they appear in the same references.
When selecting a medication/co-occurring concept pair (e.g., "adverse effects of aspirin in influenza"), WizOrder generates a fully formed MEDLINE query (e.g., "aspirin/adverse effects" AND "influenza") which is sent to the PubMed Web-based search engine to retrieve current citations, using the Web-browser built in WizOrder. In this situation, a generic search engine is used, as a simple MEDLINE query is guaranteed to return relevant references. Such an engine cannot be used for the feature described earlier, that searches for references matching the highest number of a list of MeSH terms, and iteratively adapts the search strategy based on previous results.

The same technique could also be used to provide other clinically meaningful information: co-occurrences of diagnoses and drugs with the "therapeutic use" subheading could be used to suggest potential treatments, and co-occurrences of diagnoses and procedures with the "diagnostic use" subheading could provide lists of potential investigation methods.

CONCLUSION

As a dictionary, interlingua, and knowledge source, the UMLS provides key information and linkages for the encoding of free-text and the integration of databases and knowledge bases, facilitating the seamless interfacing of clinical decision-support tools during computerized care-provider order entry.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the U.S. National Library of Medicine (5 G08 LM05443 and 1 R01 LM06226), from the Janggen-Poehn Stiftung of Saint-Gall, Switzerland, and from the Swiss National Research Fund.

References


Figure 2. Concepts are linked by semantic relations from UMLS Semantic Net. MeSH terms can also be linked based on MEDLINE co-occurrence data.
15. Miller RA, Jannback L, Giuse NB, Masarie FE. Extending the capabilities of diagnostic decision support programs through links to bibliographic searching: addition of "canned MeSH logic" to the Quick Medical Reference (QMR) program for use with Grateful Med. Proc Annu Symp Comp App Med Care. 1991; 150-5.