ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Biomedical Informatics xxx (2003) xxx–xxx

Journal of Biomedical Informatics

www.elsevier.com/locate/yjbin

Towards linking patients and clinical information: detecting UMLS concepts in e-mail[☆]

Patricia Flatley Brennan[a,*] and Alan R. Aronson[b]

[a] University of Wisconsin-Madison, 372 Mechanical Engineering, 1513 University Avenue, Madison, WI 53706, USA
[b] National Library of Medicine, Building 38A, MS 54, 8600 Rockville Pike, Bethesda, MD 20894, USA

Received 18 September 2003

Abstract

The purpose of this project is to explore the feasibility of detecting terms within the electronic messages of patients that could be used to effectively search electronic knowledge resources and bring health information resources into the hands of patients. Our team is exploring the application of the natural language processing (NLP) tools built within the Lister Hill Center at the National Library of Medicine (NLM) to the challenge of detecting relevant concepts from the Unified Medical Language System (UMLS) within the free text of lay people's electronic messages (e-mail). We obtained a sample of electronic messages sent by patients participating in a randomized field evaluation of an internet-based home care support service to the project nurse, and we subjected elements of these messages to a series of analyses using several vocabularies from the UMLS Metathesaurus and the selected NLP tools. The nursing vocabularies provide an excellent starting point for this exercise because their domain encompasses patient's responses to health challenges. In successive runs we augmented six nursing vocabularies (NANDA Nursing Diagnosis, Nursing Interventions Classification, Nursing Outcomes Classification, Home Health Classification, Omaha System, and the Patient Care Data Set) with selected sets of clinical terminologies (International Classification of Primary Care; International Classification of Primary Care- American English; Micromedex DRUGDEX; National Drug Data File; Thesaurus of Psychological Terms; WHO Adverse Drug Reaction Terminology) and then additionally with either Medical Subject Heading (MeSH) or SNOMED International terms. The best performance was obtained when the nursing vocabularies were complemented with selected clinical terminologies. These findings have implications not only for facilitating lay people's access to electronic knowledge resources but may also be of assistance in developing new tools to aid in linking free text (e.g., clinical notes) to lexically complex knowledge resources such as those emerging from the Human Genome Project.

1. Introduction

Understanding and fulfilling patient information needs is a primary nursing responsibility most often met through careful exploration of the person's chief concerns, readiness to learn, and ability to process complex content, and verbal presentation of relevant content. Advances in health informatics, including the development of standardized languages for health care, creation of internet-accessible knowledge resources and the application of electronic communication to the nurse–patient relationship, afford nurses new opportunities for facilitating patient access to health information. Search engines permit anyone with a computer network connection and a web browser to selectively explore the vast resources of literature databases, full-text journals, and informative consumer health information on the internet. However, effective use of search engines places enormous cognitive demands on patients to identify appropriate terms for exploring complex and potentially unfamiliar knowledge resources. Judicious application of existing professional vocabularies and the natural language processing (NLP) tools used to manage them may open up an alternative approach to meeting patient

[☆] Support for this work was provided by the National Library of Medicine (LM 6247), the University of Wisconsin-Madison Graduate School (Kellet Professorship) and the Lister Hill Center, National Library of Medicine, National Institutes of Health.

[*] Corresponding author. Fax: 1-608-263-5252.
E-mail addresses: kludwig@facstaff.wise.edu, pbrennan@engr.wisc.edu (P.F. Brennan), alan@nlm.nih.gov (A.R. Aronson).

1532-0464/$ - see front matter © 2003 Published by Elsevier Inc.
doi:10.1016/j.jbi.2003.09.017

information needs, one that capitalizes on the widespread acceptability of e-mail and other electronic communication mechanisms as a starting point for locating and retrieving relevant health information. The purpose of this project is to explore the feasibility of detecting terms within the electronic messages of patients that could be used to effectively search electronic knowledge resources and bring health information resources into the hands of patients.

As part of a larger program of work at the National Library of Medicine's Lister Hill Center, investigators develop and apply NLP approaches to analyze a wide range of text-based health science knowledge resources. For instance, the Indexing Initiative project investigates fully and semi-automatic methods for indexing the biomedical literature [1]. One Indexing Initiative approach employs a concept-based indexing method based on MetaMap which maps citation text to concepts in the UMLS Metathesaurus [2]. Because MetaMap employs general-purpose NLP tools, and because source vocabularies of the UMLS Metathesaurus serve as index terms for a wide range of electronic health resources of interest to laypersons (professional literature, lay literature, clinical records), these tools might be of particular value in helping identify salient concepts that appear in lay people's electronic communications. These concepts in turn could become the starting point of search and retrieval processes that directly link colloquial comments by patients to electronic knowledge resources, thus facilitating patient access to health information.

2. Background

Work in consumer health informatics over the past 15 years has focused largely on creating internet-based health information resources, communication utilities, and patient-accessible clinical records systems. Insuring clear, accurate recognition of laypersons' information needs and presenting comprehensible, appropriate responses to them form the core elements required for effective use of these resources. While most of the consumer health informatics efforts have focused on the *presentation* of health information to lay people, a smaller but critical part of these efforts attempt to *understand* patients' expressions of health care concerns in a manner that enhances their ability to access electronic knowledge resources and clinical care record systems.

Precise understanding of the concerns raised by lay people in these electronic messages usually requires interpretation by a skilled health professional. Widman and Tong [3] examined 70 messages posted to a physician-created website and determined the domain of clinical concerns (cardiology in 67 cases (96%) and related conditions and procedures in 52 cases (74%)) as well as the goals of the inquiries (diagnosis (15), therapy (48), prognosis (1), and patient education (6)). Widman and Tong also determined that psychological reassurance was a predominant reaction among those contacted afterwards. These concerns parallel those found during professional review of electronic messages received by a pediatric practice by D'Allessandro et al. [4]. They cautioned professionals responding to electronic messages to avoid personalized responses but rather refer individuals to digital library resources. Such referrals could benefit from the application of automatic indexing and mapping tools, such as those under development in the NLM's Indexing Initiative. Smith et al. [5] used one of these tools (MetaMap) to analyze human curators' culling of 504 terms reflecting findings and features from 109 questions posted to an electronic cancer information service. They found a match to UMLS concepts for over 95% of the abstracted words and phrases. Emerging NLP strategies that enhance automatic indexing may, in some cases, make it possible to eliminate the costly and time-consuming human curator participation.

Some investigators advocate creating terminology tools, including a controlled consumer vocabulary, to enhance the articulation between the phrases employed by lay people and the standardized terminologies used by health professionals. Zeng et al. [6] evaluated the terms used by lay people to search a hospital's web site and found very poor matches between the terms used by these people and those found in the hospital's website. Rather than creating consumer vocabularies, they recommended the development of extensive terminology support tools for clarifying lexical phrasing, semantic meaning, and users' mental maps during the search process. Conversely, Patrick et al. [7] determined that combining vernacular extensions to the UMLS facilitates the expression of consumer health information needs during the search process. However, strategies that emphasize creating a patient/consumer terminology then mapping from that terminology to some recognized one are labor intensive and inconsistent with the current approaches that propose linking terms to referent terminology models.

Direct searching of selected internet sites and unconstrained free-text searching of the Web generally constitutes the most frequent occasions for lay people to undertake free-text entry of terms reflecting health concepts. Exploration of this behavior provides some insights regarding how lay people best be linked to those resources that are indexed using controlled, professional vocabularies. McCray et al. [8] describe their meticulous approach to insuring that search terms entered by patients exploring the resources of ClinicalTrials.gov are recognized and result in the retrieval of relevant documents. McCray's work details management of common problems of consumer expression, such as word variants and misspellings. Eysenbach and Kohler [9] evaluated the effectiveness of internet-wide searching by lay people. Although their sources lay people frequently use search strategies that are suboptimal, most individuals are satisfied with their results. This satisfaction appears to result from the perceived ability by the individual to compare and contrast information from sites they deem relevant and of sufficient quality, and to reject those found to be irrelevant or suspicious. When approaching a search event, it is plausible that the individual user defines the task as a query and deliberately selects a small number of what they perceive to be appropriate search terms. However, using NLP tools to identify terms in the electronic messages of lay people may be advantageous because it may help extract potential search terms from a communications modality more robust and forgiving than a search term box, and makes good use of the apparent willingness of lay people to peruse suggested reference sites.

The UMLS contains six source vocabularies relevant to the domain of nursing. All nursing vocabularies have been vetted by the professional nursing association and have been subjected to rigorous evaluation for domain completeness and utility in clinical information systems [10]. Together these six vocabulary systems cover the full range of the phenomena of concern for nursing: patient problems, nursing strategies designed to manage these problems, and patient outcomes. They provide a reasonable set of reference vocabularies against which we may examine and identify the health concerns of lay people.

It is critical to remember that these vocabulary systems have particular goals and that their intent is largely to describe the nursing process and document nursing care, not to represent the concerns of lay people. However, given that the role of nursing in society is to diagnose and treat human response to disease and human development, these vocabularies address a particular part of the patient experience not addressed in other health care vocabularies. Indeed, Zielstorff et al. [11] observed that the nursing vocabularies covered concepts not found in the other vocabularies included in the UMLS. It is logical, therefore, to anticipate that the terms present in these vocabularies should provide good coverage for the concerns raised by patients in electronic mail messages to the nurse, and may be more precise and less pathology-focused than other terminologies found within the UMLS.

3. Methods

The purpose of this study was to evaluate the application of MetaMap for detecting within the electronic messages of patients the presence of terms found in the UMLS. This section details MetaMap, the source vocabularies employed in the project, the stimulus text, and the procedure followed to conduct the evaluation.

3.1. MetaMap

The MetaMap Indexing method of discovering and ranking UMLS concepts in free text consists of applying the MetaMap program to a body of text then ordering the resulting concepts using a ranking function [12]. The MetaMap program, itself, accomplishes three key functions:

1. Parse text of free-text messages from the stimulus text into simple noun phrases using the Specialist minimal commitment parser.
2. Variant generation and discovery of term candidates from the source terminologies from the UMLS.
3. Retention of the concepts associated with the best mapping terms from the candidate list using a scoring mechanism that evaluates the fit of each term from the source vocabulary to the original phrase from the stimulus text.

We constrained the MetaMap program in a few ways. While it is possible to apply MetaMap in two modes (basic, which relies on precise matching of stimulus text and source vocabulary terms, and aggressive, which tolerates extraneous words), we selected the more conservative basic mode to process the text. In all trials we evoke MetaMap in the basic mode and employed options to restrict derivational variation to noun–adjective cases, to suppress abbreviation expansion, and to ignore word order in the input text.

We conducted several trials applying MetaMap to the free text found in the electronic messages of patients, first using only the nursing vocabularies and on successive trials adding additional vocabulary sets. For each MetaMap trial, we extracted summary information such as the number of utterances and phrases processed and the number of candidate matching concepts and final mappings for each phrase. We also obtained detailed information for each phrase including concepts matched and their unique identifiers (CUIs) and semantic types, the specific vocabularies from which the matched concepts were extracted and the MetaMap score indicating how well the concepts matched the phrases.

3.2. Source vocabularies

In earlier work we employed the full UMLS vocabulary system as source vocabularies for identification of relevant terms within the electronic messages sent by patients to a nurse. This process revealed that about half of the terms parsed from patient electronic messages could be matched to one or more terms within the UMLS; however, the extent of errors, including both false positives and inappropriate mappings, suggested that greater precision could be obtained if a subset of vocabularies were employed in the exercise. The context of the messaging activity (home nursing support following discharge from surgery) might more precisely captured if the source vocabularies were restricted to those more directly relevant to home care and nursing care issues.

Therefore, in the present study we initially restricted the MetaMap application to the source vocabularies from the UMLS to those defined as nursing vocabularies and related terminologies (for example, the Thesaurus of Psychological Terms (PSY2001)).[1] Nursing vocabularies included: the North American Nursing Diagnosis Association taxonomy of nursing diagnoses [13] (NANDA99), Saba's Home Health Care Classification [14] (HHC), the Omaha System [15] (OMS), the Nursing Interventions Classification [16] (NIC 99), the Patient Care Data Set [17] (PCDS), and the Nursing Outcome Classification [18] (NOC99).

Patients experience problems and express concerns across the full spectrum of physical, psychological, and therapeutic dimensions of health and illness. Thus, for completeness, we included additional UMLS source vocabularies in our evaluation (see Table 1). These vocabularies were identified by members of our team knowledgeable about the structure of the vocabularies, the concept and semantic types included, and their organization within the Semantic Network [19]. Every concept within the UMLS is organized under one of 134 semantic types, which arrange similar terms and concepts into coherent sets (e.g., "Individual Behavior," "Signs and Symptoms"). The Semantic Network depicts relationships among semantic types in the UMLS, for example, "is a type of," "is a consequence of."

3.3. Stimulus text

For the purpose of this evaluation we obtained as source text the electronic messages sent by patients to a clinical nurse during a field evaluation of an internet-based home care post-discharge support service [20]. The HeartCare intervention provided persons recovering from Coronary Artery Bypass Graft (CABG) surgery with communication utilities and recovery coaching information accessible through a standard web browser. Following a protocol approved by the relevant Human Subjects' Committees, patients who were medically stable following the CABG procedure were randomly assigned to one of three conditions: the HeartCare intervention, audio-taped discharge coaching instructions, or usual care. Patients had access to the HeartCare intervention for a six-month period. In a 24-month randomized field investigation, patients with access to HeartCare recovered faster, and with fewer negative symptoms, than persons with access to a standardized discharge training method [21]. The messages selected for the evaluation described here came from the "E-Mail the Nurse" segment of the HeartCare intervention.

Three hundred and twenty-five sequential messages sent from patients to the project nurse were culled and anonymized according to the guidelines advanced by Sweeney et al. [22] first all surnames were extracted and replaced with subject codes unique to the current investigation. Then, references appearing in the text addressing other participants in the study were replaced with unique codes for those other participants. Finally, all elements (identifiers, headers, and message text) from the 84 message and none of the text was eliminated. An exact replica of a sample message is displayed in Fig. 1, marked with the delimiters needed for the MetaMap process. Misspellings, spacing and punctuation are as they appeared in the original message.

Table 1
Source vocabularies employed in the evaluation of patient electronic messages

North American Nursing Diagnosis Association taxonomy of nursing diagnoses (NANDA99)
Saba's Home Health Care Classification (HHC)
Omaha System (OMS)
Nursing Interventions Classification (NIC 99)
Patient Care Data Set (PCDS)
Nursing Outcome Classification (NOC99)
International Classification of Primary Care (ICPC2E)
International Classification of Primary Care- American English (ICPC2AE)
Micromedex DRUGDEX (MMX01)
National Drug Data File (NDFF01)
Thesaurus of Psychological Terms (PSY2001)
WHO Adverse Drug Reaction Terminology (WHO97)
Medical Subject Heading 2003 (MSH_2003)
SNOMED International Version 3.5 (SMNI98)

3.4. Procedure

Two hundred and forty-one messages were retained. These messages were parsed using Rindflesch's under-specified syntactic parser [23]; and the data, which included delimited phrases organized within message headers and messages, were run through the MetaMap program. To prepare the source text for analysis we arbitrarily defined the "Subject" line of each message as a "Title" and the body of each message as an "Abstract." This decision permitted us to use existing NLP tools without extensive modification. We conducted four trials, applying MetaMap in the basic mode to the stimulus text with the source vocabularies organized in the following manner:

1. Nursing Only
2. NursingPlus (the six nursing vocabularies plus International Classification of Primary Care (ICPC2E); International Classification of Primary Care- American

[1] Information on source vocabularies and natural language tools included in this project can be obtained from the UMLS Knowledge Server, http://umlsks.nlm.nih.gov/, accessed June 7, 2003.

> Subject: medications
>
> Dear Connie, I've been out of the loop for a few weeks. I had a setback with the appearance of a blood clot 2 weeks ago and was back in the hospital for a week. I was released a week ago Friday and now am on several new medications. With all these new meds, I feel nauseous almost all the time and frequently dizzy. I have a visiting nurse coming to see me 3x a week, and she monitors my blood pressure, temperature and checks my legs for possible clots. But nothing seems to help the nauseous feeling and I have little appetite. The medication I am now taking are □ I suspect the Lasix may be the culprit, since had been on it a LONG time ago and it made me nauseous, but I don't know. Do I really need to be on all of these now? I take alot of them at the same time (meal time), but should I change this and stagger them? What order should I take them, or are there alternatives to this medication for now? Any advise you could give me before I go back to see my internist on Tuesday would be helpful, then I could discuss it with him again. I see the cardiologist on Thursday and hope to be cleared to start cardiac rehab after that. Right now, however, it is slow going and discouraging. Thanks, Bill

Fig. 1. Sample message.

English (ICPC2AE); Micromedex DRUGDEX (MMX01); National Drug Data File (NDFF01); Thesaurus of Psychological Terms (PSY2001); WHO Adverse Drug Reaction Terminology (WHO97)).
3. NursingPlus and Medical Subject Headings (MeSH)
4. NursingPlus and SNOMED International 3.5 (SNMI)

We conducted our evaluation on two levels: first, we evaluated the entire stimulus text as a unit successively using MetaMap to apply the four vocabulary sets. Next we selected a single message and examined the results of the four mapping trials for it.

Quantitative measures of vocabulary coverage provide necessary but not sufficient evidence for appraising the adequacy of a source vocabulary to capture the terms employed by lay people in their electronic messages. It is also instructive to examine the source text in detail to explore the actual matches and to pay particular attention to the types of mismatches (errors) that occur in the process. For this purpose, we selected a message with a sufficiently dense stimulus text (count of words, diversity of topics) to illustrate as broad a range as possible of terms. Table 2 summarizes the evaluation of the message depicted in Fig. 1 across all four vocabularies.

This single message yielded 174 phrases. Performance of the vocabularies on the four trials paralleled that observed in the trials of the entire source test in that the mean number of matches per phrase remains close to 1 and increases as vocabularies are added. What is instructive here is the change in the number of errors in matching that occurred during the mapping process. We defined errors based upon review of the source text by the primary author. Three types of errors were found:

1. Recognition
   - 1.1. Stimulus terms are parsed in an overly granular manner
     - 1.1.1. *the words "feeling nauseous" are parsed as two separate terms*
2. Inappropriate terms, concepts or semantic types selected from the source vocabularies
   - 2.1. Concept mapping is nonsensical
     - 2.1.1. *The word "I've" is mapped to the term VAL (Semantic Type = Disease or Syndrome)*
   - 2.2. One or more of the Semantic types is inappropriate for the context
     - 2.2.1. *The word "back" is mapped to the term "back" but the semantic type indicated was "Body Location or Region"*
3. Matching
   - 3.1. Polysemy—word that has more than one meaning
     - 3.1.1. *The word "monitor" can refer both to a Health Care Activity and a Medical Device*
   - 3.2. Semantic types hold radically different meanings in different vocabularies
     - 3.2.1. *The word has more than one meaning, e.g., "right" within SNOMED has the Semantic Type "Spatial Concept" but within MeSH is a "Qualitative Concept"*

Table 2
Analysis of a single message (174 phrases parsed)

	Nursing Only	NursingPlus	NursingPlus + MeSH	NursingPlus + SNOMED
Candidates concepts	15	54	85	114
Mapped concepts	13	42	57	70
Phrases w/ one or more maps	12	43	50	57
Mean concepts/phrase	1.08	.98	1.14	1.22
Errors	3	23	37	39

Table 3
Results of the four MetaMap mapping on the full stimulus text

	Nursing Only	NursingPlus	NursingPlus + MeSH	NursingPlus + SNOMED
Candidates concepts	1016	3734	5786	7366
Mapped concepts	948	3094	4439	5078
Phrases w/ one or more maps	871	2863	3995	4383
Mean concepts/phrase	1.09 (0.28)	1.08 (0.30)	1.11 (0.35)	1.16 (0.38)

4. Results

The stimulus text yielded 241 messages. Thematic content of the messages addressed symptom management, activities of every day living, and logistics of study participation. Application of the parsing process yielded 15,326 distinct phrases, i.e., terms. The MetaMap program nominated matches to these terms from the concepts in the source vocabularies in a string matching process (see [2] for details on this approach). It is possible to have more than one term from the stimulus document matched to a single concept in the source vocabularies. The summary of results for all four trials is presented in Table 3.

Candidate concepts include all terms from any included source vocabulary deemed to be an eligible match to the parsed phrase (For the Nursing Only run, this number is 1016). Mapped concepts include only those retained by the MetaMap evaluation process, 948 for the Nursing Only run. Matches were found for only 871 of the total 15,326 parsed phrases. Thus, for the Nursing Only run, the mean number of matches per phrase was 1.09 (s.d.0.28). A mean number of matches per phrase close to 1 is desired, as this would indicate a precise and unambiguous fit of the concept to the phrase.

Table 4 summarizes the vocabularies, the number of concepts and semantic types in each vocabulary, and the coverage provided by each vocabulary for the entire stimulus text. The number of terms from the vocabularies used in the mapping exercises is generally greater than the total number of terms in the vocabulary itself, indicating that some vocabulary terms are matched to more than one stimulus phrase. Two nursing vocabularies, the Omaha System and the Nursing Outcomes Classification, provided the best source of concepts for matching to the stimulus text. In general, as the number and complexity of vocabularies included in each trial of the mapping exercise increased, the coverage of source text phrases also increased.

4.1. Discussion

The application of the MetaMap process to decode UMLS concepts from the electronic messages of lay people yielded promising results. Each trial detected important concepts present in the messages, but each trial also exposed significant limitations. While the Nursing vocabularies alone provided the best ratio of terms mapped to errors (13/3), the addition of salient clinical terminologies (NursingPlus) yields the best results, balancing the coverage of the terminologies (42/

Table 4
Description of the vocabularies and the number of times terms from that vocabulary were used in a match in each of the four trials

Vocabulary descriptions			# Matched terms			
Vocablary	# Terms	Semantic types	Nursing Only	Nursing Plus	Nursing Plus + MeSH	Nursing Plus + SNOMED
NANDA	169	12	272	258	256	251
NIC	10187	21	29	29	28	26
NOC	3056	31	468	426	415	398
HHC	335	26	112	98	96	91
OMS	539	48	395	351	336	335
PCDS	2229	25	105	87	85	82
PSY 2001	7671	119		2747	2691	2641
WHO 97	3831	37		258	260	231
ICPC2AE	210	14		8	8	8
ICPC2E	3757	42		244	242	234
MMS 01	11536	36		204	161	182
NDDF01	20088	47		241	198	218
MESH	516793	134			3519	
NOMED	164179	131				4245

174) with the ratio of concepts mapped to errors generated (42/23). The addition of the vocabularies in NursingPlus appears to offer benefit both because of the increase in concept coverage as well as the increase in the number of Semantic Types included in these vocabularies.

The original motivation of this study was to explore ways to use the nursing vocabularies for interpreting the concerns of lay people. Our study then should be considered a success in that it demonstrates that the Nursing Only vocabularies do provide an accurate, if incomplete, representation of the terms patient use in their electronic mail messages. When deemed correct, the terms from the source vocabularies provided a reasonable fit to the phrase from the stimulus text. For example, the phrase "the nauseous feeling" mapped to the term "nausea" extracted from several vocabularies from the Nursing Outcomes Classification and the NANDA Nursing Diagnoses, and the phrase "little appetite" mapped to the NOC term "Appetite."

The study was narrow in scope, and in no way does this work represent an attempt to interpret free text messages of lay people or generate automatic responses to complex, unstructured queries. It is critical to remember that the goal was to simply determine if concepts the standardized vocabularies used for professional purposes could be found among the free text messages of lay people. This well-circumscribed goal was accomplished, and our results show that it is possible to detect concepts from standardized vocabularies, including the nursing vocabularies, in the free-text of lay people. However, these results also show that large amounts of the free text messages of lay people do not include concepts from the standardized vocabularies present in the UMLS and that the mapping of these source terms to the stimulus phrase remains imprecise at best.

It is appropriate to conclude that in subsequent studies, the NursingPlus vocabulary set may provide the best starting point. Several caveats are in order; however, first, one must be mindful of the expectation of vocabulary coverage available from a single vocabulary. Without considering errors, the best performing vocabulary still only provided matches for one third of the phrases identified. However, given the wide range of topics found in even a single electronic message, it is unlikely that as large a number of phrases parsed would be recognized by specialized terminologies as is found in evaluation of the controlled text of professionals. This result does suggests that future work applying the NLM NLP tools to the free text of lay persons may require modifications in the parsing process so that idioms and other colloquial written speech employed by lay people be appropriately recognized.

Second, it is important to conceptualize the process of decoding patient electronic messages as a multi-step process. This study addressed the first step, determining whether it is possible to identify plausible concepts from the UMLS. Additional post-processing activities, such as selection of the best search terms from a list of mapped candidates, might be in order, and require the generation of rules based on additional information such as the identity of the user. It is possible that post-processing of the list of mapped concepts to generate a reduced list of potential search terms will compensate for some of the errors evidenced here.

The application of the NLP tools, with the vocabularies of the UMLS restricted to the nursing vocabularies and related terminologies, provided greater precision with fewer errors than did our earlier application of the tools using the full UMLS. Therefore, our approach offers support for the application of the tools to consumer health issues; but it recommends that a restricted set of vocabularies be used during the application of the tools to unstructured text.

This application of the NLM tools is non-standard, and represents an extension not envisioned by the developers. We chose to restrict our processing of electronic messages to the constraints imposed by the tools, including delegation of the subject line to the expected "Title" designation and the message body to the "Abstract." Both the titling and abstracting of professional articles follows expectations of systematic thinking that may not be present in the construction of electronic messages by lay people. The primary compromise here is the presumption of coherence between the two. Multiple themes are present in almost every electronic message, thus suggesting that electronic messages consist of lists of almost unrelated topics. However, because both the parser and the source vocabularies target parsing of terms from source text independent of context, this modification is logical and acceptable within bounds.

The tools show some promise in identifying relevant professional health care vocabularies in the terms found in the free text of lay people employed in electronic messages. Additional work is required to scale the procedure employed here for routine use in the analysis of electronic messages. Greater attention to mapping accuracy, error determination, and management of errors is needed before general-purpose use of this approach can be advocated. In summary, the development of tools that will assist health professionals and lay people in identifying the health concepts present in the free-text electronic messages of lay people holds great promise, but awaits future pre-processing and post-processing strategies.

4.2. Future applications

Patients are more likely to use electronic resources recommended by their health professionals, and refinements in the procedures described in this study could facilitate health professionals' recommendations without a concomitant investment in human curation. While our approach is designed to complement both human curation and automatic interpretation, it does not replace either. Nonetheless, it offers significant advantages to both of these in that it could be invoked deliberately as a plug-in utility, allowing the sender to mark up and gain access to health information prior to professional interpretation, which may in turn provide answers to some questions immediately or result in more precise and informative messages to professionals, thus making better use of the scarce resource of health professionals.

The approach employed here to detect concepts from recognized source vocabularies in the free text of patients may provide a model for other situations in which the need exists to create linkages between free text descriptions of clinical or biomedical phenomena and electronic knowledge bases indexed by specialized vocabularies. For example, this process may help detect in the clinical notes of primary care clinicians caring for patients with heritable diseases the presence of concepts that are also found in taxonomies of genetic diseases.

5. Conclusion

The NLP tools of the NLM show promise for their utility in identifying words and phrases in the free text of lay people. Greater precision and coverage, and fewer errors, were found through applying the MetaMap process employing the NursingPlus vocabularies. Thus, while it is possible to detect within the written text of lay people terms found in standardized nursing vocabularies, full utility awaits more efficient pre-processing, that insures detection of the perhaps-unusual expressions of lay people and post-processing, which will refine the precision of mapping concepts to the vernacular of lay people.

References

[1] http://ii.nlm.nih.gov Accessed April 7, 2003.
[2] Aronson AR. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. Proc AMIA Symp 2001:17–21.
[3] Widman LE, Tong DA. Requests for medical advice from patients and families to health care providers who publish on the World Wide Web. Arch Intern Med 1997;157(2):209–12.
[4] D'Alessandro DM, D'Alessandro MP, Colbert SI. A proposed solution for addressing the challenge of patient cries for help through an analysis of unsolicited electronic mail. Pediatrics 2000;105(6):E74.
[5] Smith CA, Stavri P, Chapman WW. In their own words? A terminological analysis of e-mails to a cancer information service. Proc AMIA Symp 2002:697–701.
[6] Zeng Q, Kogan S, Ash N, Greenes RA, Boxwala AA. Characteristics of consumer technology for health information retrieval. Methods Inf Med 2002:289–98.
[7] Patrick TB, Monga HK, Sievert ME, Houston Hall J, Longo DR. Evaluation of controlled vocabulary resources for development of a consumer entry vocabulary for diabetes. J Med Internet Res 2001;3(3):E24.
[8] McCray AT, Dorfman E, Ripple A, et al. Usability issues in developing a Web-based consumer health site. Proc AMIA Symp 2000:556–60.
[9] Eysenbach G, Kohler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 324:573–7.
[10] Bakken (Henry) S, Warren JJ, Lange L, Button P. A review of major nursing vocabularies and the extent to which they have the characteristics required for implementation in computer-based systems. JAMIA 1998;5(4):321–8.
[11] Zielstorff RD, Cimino C, Barnett GO, et al. Representation of nursing terminology in the UMLS metathesaurus: a pilot study. Proc Annu Symp Comput Appl Med Care 1993.
[12] Aronson AR, Olivier B, Chang F, et al. The NLM indexing initiative. Proc AMIA Symp 2000:17–21.
[13] NANDA. Nursing diagnoses: definitions and classification 2001–2002. North American Nursing Diagnosis Association, Philadelphia: 2001.
[14] Saba VK. Home Health Care Classification (HHCC): Nursing Diagnosis and Nursing Intervention. HHCC Web site. Available from: www.sabacare.com, accessed April 7, 2003.
[15] Martin KS, Scheet NJ. The Omaha system: applications for community health nursing. Philadelphia: WB Saunders; 1992.
[16] McCloskey Joanne C, Bulechek Gloria M, editors. Nursing interventions classification (NIC): iowa intervention project. St. Louis (MO): Mosby Year Book; 1999.
[17] Ozbolt JG, Russo M, Stultz MP. Validity and reliability of standard terms and codes for patient care data. Proc Annu Symp Comput Appl Med Care 1995:37–41.
[18] Johnson M, Maas M. The nursing outcomes classification. J Nurs Care Qual 1998;12(5):9–20.
[19] http://umlsks.nlm.nih.gov, accessed June 7, 2003.
[20] Brennan PF, Moore SM, Bjornsdottir G, Jones J, Visovsky C, Rogers M. HeartCare: an Internet-based information and support system for patient home recovery after coronary artery bypass graft (CABG) surgery. J Adv Nurs 2001;35(5):699–708.
[21] Moore SM, Brennan PF, O'Brien RA. HeartCare: Early Results. (Abstract) AHA Nov 01.
[22] Sweeney L. Three computational systems for disclosing medical data in the year 1999. Medinfo 1998;9(Pt 2):1124–9.
[23] Rindflesch TC, Rajan JV, Hunter L. Extracting molecular binding relationships from biomedical text. Appl Nat Lang Process 2000:188–95.