Identifying Consumer-Friendly Display (CFD) Names for Health Concepts

Authors: Qing T. Zeng[a], Tony Tse[b], Jon Crowell[a], Guy Divita[bc], Laura Roth[c], Allen C. Browne[b]

Affiliations:

[a]Decision Systems Group
Brigham and Women's Hospital
Harvard Medical School
Boston, MA

[b]Cognitive Science Branch
Lister Hill National Center for Biomedical Communications
National Library of Medicine
National Institutes of Health
Department of Health and Human Services
Bethesda, MD

[c]Management Systems Designers, Inc.
Fairfax, VA

ABSTRACT

We have developed a systematic methodology using corpus-based text analysis followed by human review to assign "consumer-friendly display (CFD) names" to medical concepts from the National Library of Medicine (NLM) Unified Medical Language System® (UMLS®) Metathesaurus®. Using NLM MedlinePlus® queries as a corpus of consumer expressions and a collaborative Web-based tool to facilitate review, we analyzed 425 frequently occurring concepts. As a preliminary test of our method, we evaluated 34 analyzed concepts and their CFD names, using a questionnaire modeled on standard reading assessments. The initial results that consumers (n=10) are more likely to understand and recognize CFD names than alternate labels suggest that the approach is useful in the development of consumer health vocabularies for displaying understandable health information.

INTRODUCTION

Consumer health informatics has the *potential* to transform the healthcare system [1]. Researchers are investigating a range of consumer health applications, including online access to medical records [2,3], direct messaging between patients and clinicians [4], medical data entry by patients [5], and patient decision-support tools [6]. Despite the increasing availability of these tools, consumers continue to have difficulty finding, understanding, and applying the health information provided [7].

Health literacy is a significant barrier to accessing health information: "Over 300 studies, conducted over three decades and assessing various health-related materials, such as informed con-

sent forms and medication package inserts, have found that a mismatch exists between the reading levels of the materials and the reading skills of the intended audience" [8]. While educating consumers and improving providers' communication skills are long-term solutions, other strategies are necessary to address current consumer health information needs.

One approach is to link common health-related language to professional medical concepts through consumer health vocabularies (CHVs), reviewed by Zielstorff [9]. Unlike traditional terminologies that are prescribed from expert and domain sources (e.g., literary warrant), CHVs are *experiential* and descriptive of everyday usage and needs by consumers in making sense of health-related topics and issues (e.g., use warrant).

In this paper, we describe a systematic approach for CHV development that combines corpus-based text analysis and human review. We recognize that CHVs serve multiple purposes (Table 1) and that each purpose may require a different "consumer-friendly" label. For example, the display name "Human Immunodeficiency Virus (HIV)" clearly and unambiguously refers to the causative agent of AIDS and invokes a particular concept. It is likely that such a string would be a good display name, as consumers who read that string and are familiar with either form, "human immunodeficiency virus" or "HIV" will recognize the underlying concept. However, the same string would not an optimal term for information retrieval (IR), as illustrated by results on Google (accessed on 3/18/05):

- Human Immunodeficiency Virus (HIV):   9.9 million hits
- Human Immunodeficiency Virus: 1.6 million
- HIV: 36.9 million

This study focuses on an important CHV subtask: identifying "consumer-friendly display" (CFD) names. That is, expressions (i.e., words or phrases) describing medical concepts likely to be recognized by most consumers. A preliminary evaluation of CFD names is reported, independent of context. Note that selection of consumer-friendly names for other tasks is not addressed here.

CHV Subtask	Description
Display (CFD$_i$), context independent	"Preferred name" for a medical concept
Display (CFD$_d$), context dependent	Name for a medical concept that will most likely be familiar to a consumer within a larger context (e.g., embedded in text)
Information Retrieval (CFR)	Name(s) familiar to a consumer that will optimize searching, providing relevant IR results
Mapping (CFM)	Name(s) most likely to be useful for extracting consumer expressions from natural language text

Table 1. CHV subtasks and requirements for "consumer friendly" names

BACKGROUND

While differences between the language of laypersons and professionals in the medical domain have been studied (e.g., [10-12]), we recently embarked on the development of an open-source

first-generation CHV[1] to bridge the vocabulary gap. As a first step, a method for systematically identifying CFD names for medical concepts was developed.

The medical domain is intimidating to many consumers, particularly due to its specialized terminology and the potentially serious health consequences resulting from miscommunication. It is estimated that nearly 70% of medical terms are derived from Greek or Latin [13]. Thus, it is not surprising that even highly literate consumers may stumble over professional medical jargon full of arcane Greco-Latin terms representing complex biomedical and clinical concepts. Even highly literate consumers may stumble over medical jargon. Luckily, many health-related concepts may be represented by terms that are more familiar to lay people, such as "external ear" (*pinna*) and "burp" (*eructation*). Thus, identifying and using CFD names may facilitate communication, such as:

- System applications that display familiar names for health concepts to consumers
- Translation of less understandable names into familiar ones within health text intended for consumers.

Identifying CFD names is a variation of a long-recognized vocabulary problem. Variability in names of objects or concepts is common in everyday language and influences factors such as personal experience, knowledge, and membership in discourse groups [14]. For example, one study reported a less than 20% likelihood of two people using the same name for a concept in domains as diverse as common objects, ingredients in recipes, items found in classified ads, and text-processing commands [15].

From a "display" perspective, understanding text is impeded if a word or phrase used by a provider or author of health text is not common, has a different connotation, or is unfamiliar to the intended readers. On the other hand, from a "literacy" perspective, text comprehension is impeded by unfamiliar words or phrases or those having distinct connotations. The goal of the CFD name task is to find a single well-known, unambiguous label for each medical concept.

In fact, health literacy experts have created substitute word lists of CFD names [16]. However, such lists typically contain only several hundred names, serving as examples and teaching aids. A few studies have tested consumers directly regarding their knowledge of medical terms (e.g., [17,18]) or common colloquialisms (including slang) and malapropisms used by consumers that providers may not know [19]. While these results demonstrate specific differences between consumer and provider language, the methods were not intended to be systematic or scalable.

The most direct and authentic way to identify CFD names is to ask a representative sample of consumers to review lists of expressions and record recognition accuracy and frequency. However, given the diversity among healthcare consumers and the vast number and range of health-related concepts, the resources required are likely to be prohibitive.

Text analysis provides a feasible alternative. In the study we report here, frequently occurring expressions submitted to a consumer health information site as queries were collected. We made an underlying assumption that search frequency correlates with recall and recognition. That is, the more frequently an expression occurs in submitted search strings, the greater the likelihood

---

[1] See http://consumerhealthvocab.org/

that typical users of that Web site would be familiar with the expression. Because frequency alone is insufficient for identifying CFD names (e.g., due to word sense ambiguity), human review is also required.

METHODS

We developed a two-step approach to investigate CFD names. In the first step, we mapped frequently used consumer expressions to the Unified Medical Language System® (UMLS®) Metathesaurus® (2004AA) using lexical processes. During the second step, we reviewed expressions matched to the most common UMLS concepts, discussed candidate names, and voted on CFD names. To evaluate the CFD names, we used a questionnaire based on the Test of Functional Health Literacy Assessment (TOFHLA) [20].

1. Candidate Name Generation

We generated candidate CFD names using automated text analysis and mapping procedures. For the text analysis, we used a corpus of all queries submitted to NLM MedlinePlus® [21] over a 12-month period (October 2002-03). MedlinePlus query logs represent one of the largest and most diverse corpora of consumer-generated, health-related expressions.

The logs were preprocessed to filter out (1) non-English terms, using the UMLS SPECIALIST® lexicon to identify words in English; (2) multiple queries from the same IP address, which indicates machine-generated strings; and (3) redundant, identical strings submitted from the same IP within 5 minutes. We then mapped the remaining queries to the UMLS using lexical processes similar to MMTx[2], including removal of non-alphanumeric characters, stemming, normalization, and truncation.[3]

All expressions mapped to a UMLS concept were considered to be CFD name candidates (except those manually identified as improper mappings, resulting from aggressive stemming and normalization). Synonyms listed in the Metathesaurus but not found in the log data were not considered candidates, except for UMLS preferred terms. Thus, only mapped expressions from MedlinePlus queries and UMLS preferred terms were manually reviewed.

Authentic, spontaneous consumer utterances (e.g., transcripts from patient medical histories) would provide an ideal source of candidate CFD names. Because such source material is difficult to obtain, we used MedlinePlus query logs, recognizing the limitations (e.g., queries submitted by professionals, expressions copied from professional or media sources).

2. Collaborative Name Review

We collaboratively reviewed candidate names to select CFD names. An ideal CFD name satisfies all three criteria: (1) usefulness to consumers (as indicated by frequency of usage); (2) clarity (excludes colloquialisms and slang); and (3) readability. As discussed, although usage frequency is an indicator of familiarity, human review is essential to determine face validity. During multiple rounds of collaborative review, we developed a process for selecting CFD names, as described below.

---

[2] http://MMTx.nlm.nih.gov/
[3] Unmapped query terms are being used for another aspect of CHV development not discussed in this paper.

a. <u>Concept Usefulness</u>. Some concepts were determined to be too vague or obscure for a CHV. For example, "testing" (C0039593) is vague and "Cancer Genus" (C0998265) is obscure. No CFD name was considered for such stop concepts.

b. <u>Concepts as Modifiers</u>. The review of modifier concepts as a class, such as "acute" (C0205178) and "Red color" (C0332575), was deferred because the semantics of modifiers are frequently context-sensitive and difficult to define (e.g., "redness" as a normal or pathological state; degree of "redness").

c. <u>Term Validity</u>. Some expressions were too vague or ambiguous for a CHV (e.g., "of," "in"). No stop terms were considered as candidate CFD names.

d. <u>Mapping Appropriateness</u>. Expressions resulting from improper lexical mappings (e.g., described by Divita et al [22]) were disqualified as candidate CFD names. For example, the mapping of the expression "depression" to the concept "Cancer patients and suicide and depression" (C0812393) was deemed incorrect; these was no evidence from the log data that consumers actually used the expression to represent the concept. Similarly, "Liver" for "Liver brand of Vitamin B 12" (C0721399) were considered to be inappropriately mapped. Note that several methods were used to determine the "intended meaning" of expressions: reviewing the context of queries containing the expressions,, including referring to general and medical dictionaries intended for laypersons, and searching the Web for common usage patterns.

e. <u>Assignment of CFD Names</u>. We attempted to select expressions that unambiguously refer to UMLS concepts and are "familiar" to or easily understood by consumers (i.e., "consumer friendly"). That is, exposure to a preferred display name should trigger unambiguous and appropriate mental associations with the underlying medical concept.

For example, "Cancer" is the CFD name for "Malignant Neoplasms" (C0006826) because it is (1) the most frequently used expression mapped to the concept; (2) semantically unambiguous; and (3) a common word. In contrast, although "Medicine" occurs frequently, it potentially refers to several UMLS concepts (e.g., "Pharmaceutical Preparations" (C0013227), "Science of Medicine" (C0025118)). In all cases, reviewers had to apply personal judgment.

f. <u>Creation of CFD Names</u>. If no candidate names were deemed appropriate, CFD names were proposed by reviewers. For example, since "Diethylstilbestrol" does not occur frequently and its acronym "DES" does, but is ambiguous and not highly readable, the CFD name "Diethylstilbestrol (DES)" was created.

Six reviewers identified CFD names independently. Disagreements were resolved through discussion. Because the concepts and expressions that occur at a higher frequency are indicators of greater utility to consumers, the review process began with the highest frequency concepts. Only candidate names that occurred at least 10 times in the log data were reviewed.

Since the process involved multiple participants from geographically distributed locations and consisted of multiple rounds of review, we developed a Web-based computer-supported coopera-

tive work (CSWC) application, **VocabTool** (Figure 1). In this study, reviewers used the Preferred Names module to examine concepts, candidate terms, and log data—including contextual information through the UMLS Semantic Navigator—and enter detailed comments that could be accessed by other reviewers. The tool also allowed reviewers to generate reports on the fly to track differences in opinions and to record final CFD names after consensus was reached.
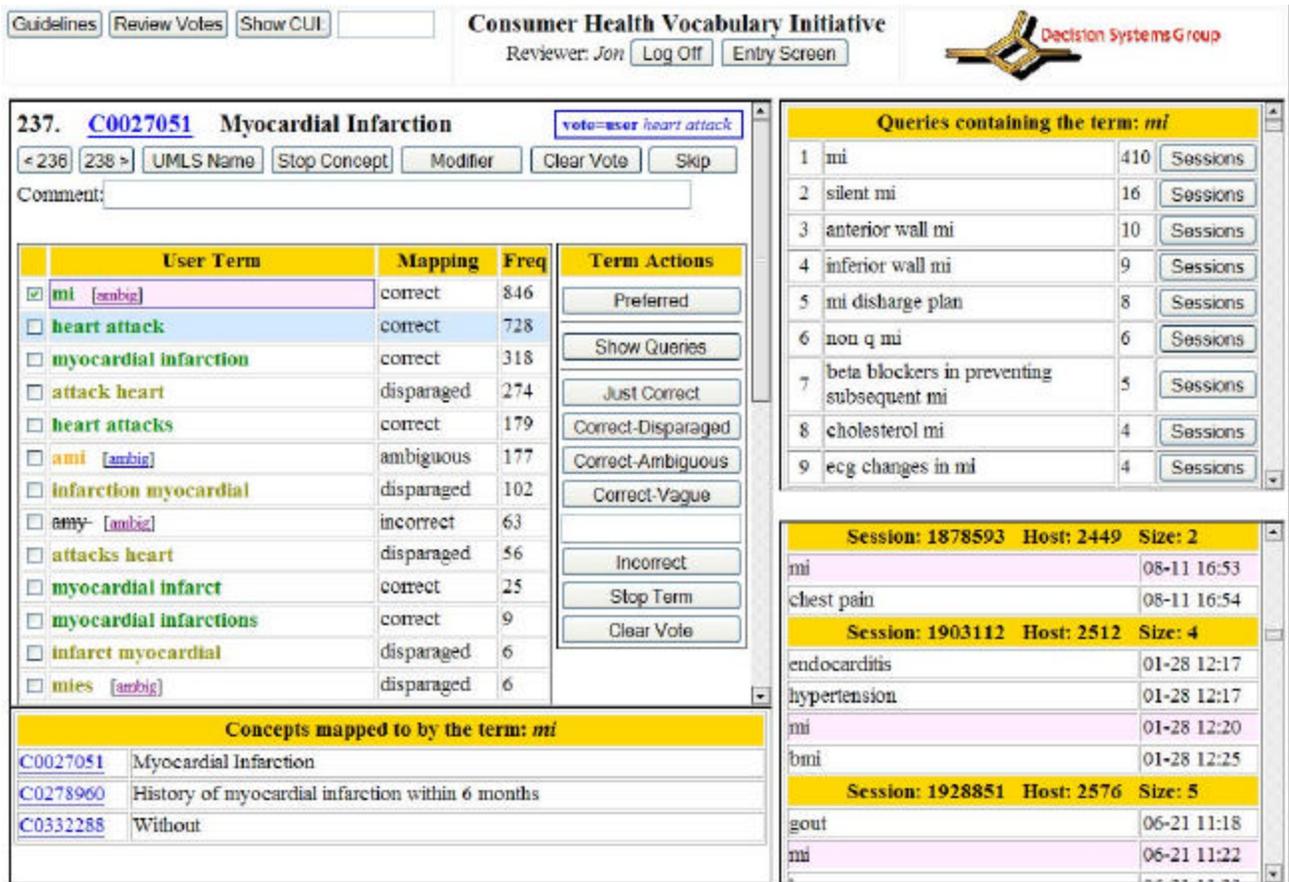

Figure 1. Partial screen shot of VocabTool application

The VocabTool also contains two additional modules (Figure 2):
- N-grams: allows users to set parameters for retrieving n-grams extracted from a consumer health corpus
- Terms/Concepts/Relations: provides users with the ability to create terms, concepts, and relationships that should be represented in a CHV, but not in the UMLS, for example

Thus, VocabTool is designed to be a full-featured CHV development environment to facilitate human review.

Figure 2. Partial screen shots displaying VocabTool's three modules

3. Evaluation of CFD Names

A preliminary evaluation study with 10 participants was conducted to determine whether the CFD names identified are more comprehensible than alternate names. We devised a questionnaire modeled after the reading comprehension part of the TOFHLA [20], a popular health literacy test among researchers. Our questionnaire contains 34 fill-in-the-blank questions, each with four multiple-choice selections: an answer and three distractors (Figure 3).

CFD Version:

Lung disease might cause _____.

     A. coughing and difficulty breathing

     B. pain in eyes and ears

     C. sudden changes in one's mood

     D. frequent indigestion

Non-CFD Version
(Same multiple choices as the CFD version):

Pulmonary disease might cause _____.

Figure 3. Sample evaluation question (2 versions)

Each question, designed to test a person's ability to understand a health concept, has two versions—one using the CFD name of a concept; the other using either the UMLS preferred term or the most frequently used alternate name (other than a lexical variant of the CFD name). The 34 concepts were selected semi-randomly from the entire set of manually reviewed concepts. All authors participated in the construction of the questions and distractors.

Participants (n=10; non-clinician, =18 years old, English speaking) were recruited from the lobbies of the Brigham and Women's Hospital. Each was provided with a copy of the questionnaire on paper. Half received a version in which the even-numbered questions contained CFD names; for the other half, the odd-numbered questions containing the CFD names.

Responses were scored as follows: +1 point for a correct answer; -1 point for an incorrect answer; and 0 points for no answer. A paired t-test was used for the hypothesis that the mean score on CFD questions was greater than that on non-CFD questions.

RESULTS

The study results are presented in three sections parallel to the description of the methodology: each phase of the two-step approach and the evaluation.

1. Candidate Name Generation

In all, 12.5 million queries were processed and the resulting consumer expressions mapped to 96,029 unique UMLS concepts. Of these, the most frequently mapped consumer expression *dif-*

fered from the corresponding UMLS preferred name for 42,619 concepts (44%). As described in the Methods section, all of these are UMLS synonyms or their lexical variants. Through this process, total of 195,140 CFD name candidates were obtained or 2.0 candidates per concept on average. As expected, the more frequently used concepts tended to have more candidate names: an average of 6.4 candidates per concept was identified among the top 1,000 concepts.

2. Collaborative Name Review

We manually reviewed 425 concepts (including stop concepts and modifier concepts) and assigned CFD names to 294 (70%). Although the concepts reviewed account for only a fraction of the number of unique mapped-to concepts, they represent 35% of all concepts mapped to expressions from the log data set. To ensure consistency, all six authors reviewed the first 224 concepts (training set). They initially reviewed 102 of these concepts as a group and reached consensus on coding policies. Independent coding of 122 concepts using the preliminary coding policies resulted in 48% complete agreement; 30% majority agreement (similar coding among at least 4 of 6 authors); and 21% lacking majority agreement.

In the test set, each of 201 concepts was reviewed by two authors independently, following the final coding guidelines described in the Methodology section. A third reviewer acted as a "tiebreaker" when required. Overall, there was 69% total agreement and 31% required a third reviewer. Nearly complete agreement was reached following discussion.

Of the 296 concepts with an identified CFD, UMLS preferred terms were selected as the CFD names (e.g., "infant" (C0021270)) for 55% of reviewed concepts. Another consumer expression mapped to the concept (i.e., a UMLS synonym) was deemed to be a CFD name for 30% of concepts (e.g., "drug" for "Pharmaceutical Preparations" (C0013227)). Finally, CFD names created using the naming policy account for 14% of reviewed concepts (e.g., "human immunodeficiency virus (HIV)" for "HIV" (C0019682)).

3. Evaluation of CFD Names

In a preliminary evaluation, a total of 10 volunteers completed the questionnaire. Among the 4 women and 6 men, the average education attained was high school and mean age was 44 years.

On average, responses were provided for 30 questions out of the 34 total (88%). In particular, participants attempted to answer more questions containing CFD names than ones with alternate names. Every subject scored higher on the CFD questions. Overall, the mean score for CFD questions was 15.4, compared with 6.0 for non-CFD questions. Thus, even in this small sample, a statistically significantly difference (p<0.01) was detected: subjects scored better on CFD questions than non-CFD questions.

Among the 34 questions, several CFD/non-CFD name pairs were recognizable by consumers, such as *Infants/Babies* and *Fracture/Broken Bone*. For other pairs, the CFD name was clearly much more familiar to consumers than the non-CFD, professional label, such as *Rash/Exanthema* and *Itching/Pruritus*.

DISCUSSION

We used a combined text-analysis and collaborative human-review approach to identify CFD names for commonly used health concepts. We designed and tested an approach to evaluate whether consumers found the CFD names we identified to be more comprehensible than corresponding alternate names. Although some health terms are known to be more comprehensible for the lay audiences than others (e.g., "stroke" and "cerebrovascular accident"), we are unaware of any previous efforts to identify and evaluate CFD names for medical concepts systematically. We believe that CFD names will improve the comprehensibility of health concepts and, ultimately, benefit health communication.

We found text analysis and manual review to be critical methods. Corpus-based text analysis not only provides candidate names and frequency information, but also helps reviewers "interpret" the meaning intended by consumers (semantics). Although manual review is time-consuming, human judgment and world knowledge are essential for CHV development and CFD name identification. For example, the expression "depression" was the most frequently mapped-to name for five UMLS concepts, including "Mental Depression" (C0011570), "Depressive disorder" (C0011581), and "Cancer patients and suicide and depression" (C0812393). Only human review, with support of authentic contextual cues from user queries containing the expression, could determine:

1. Appropriateness of the mappings in a CHV context (e.g., mapping to "cancer patients and suicide and depression" (C0812393) was deemed incorrect) and
2. Likelihood that a lay expressions represents a CFD name for any of these concepts (e.g., "depression" was considered to be a vague term, and as such, could serve as a CFD name for several related medical concepts).

As an initial attempt to identify CFD names, our work has not addressed many nuances. In order to simplify our task for this initial study, we only considered a single CFD name per concept and did not measure the comprehensibility of concepts or expressions as a continuous variable. While we treated all consumers as a single population, we recognize that many subgroups exist (e.g., non-native English speakers, differences in cultural, educational, or economic experiences) and influence health literacy and familiarity with health vocabulary (e.g., slang). The use of CFD names and professional labels is not mutually exclusive. Not only are technical terms appropriate in certain settings, they are required to educate lay persons [23]. That is, the CFD name can serve as an "entry point" to the medical term/concept. In addition, the notion of CFD name has no impact on people lacking any knowledge of a concept.

We recognize that manual review of all unique concepts from this single source (over 96,000 concepts), let alone text sources representing other discourse groups, is not feasible or scaleable. However, our strategy is to begin with concepts that have the highest usage frequencies. For example, in this study we reviewed 425 UMLS concepts that account for 35% of total mapped-to concepts in the log data set. Our goal is to review the top 1,000 mapped-to concepts which account for ~50% usage. Thereafter, review of the next 4,000 most frequent mapped-to concepts might be achievable in months, thereby accounting for ~80% usage for the top 5,000).. We also realize that this approach neglects consumer expressions that fail to map to UMLS concepts

automatically or for which no comparable UMLS concepts exist (as found in previous studies). Those issues require different approaches, which we have begun studying.

CONCLUSIONS

The work reported here is part of an effort to develop a first-generation open source CHV. We developed a two-step approach that combines text analysis and human review to identify CFD names for health-related concepts. The approach was supported through a preliminary evaluation, which showed statistically significantly better comprehension scores of CFD names compared to alternate labels.

ACKNOWLEDGMENTS

REFERENCES

1 Eysenbach G. Consumer health informatics. BMJ. 2000 Jun 24;320:1713-6.

2 Masys D, Baker D, Butros A, Cowles KE. Giving patients access to their medical records via the internet: the PCASSO experience. J Am Med Inform Assoc 2002;9(2):181-91.

3 Cimino JJ, Patel VL, Kushniruk AW. The patient clinical information system (PatCIS): technical solutions for and experience with giving patients access to their electronic medical records. Int J Med Inf 2002;68(1-3):113-27.

4 Delbanco T, Sands DZ. Electrons in flight— e-mail between doctors and patients. N Engl J Med 2004;350(17):1705-7.

5 Wald JS, Bates DW, Middleton B. A Patient-controlled Journal for an Electronic Medical Record: Issues and Challenges. Medinfo;2004:1166-72.

6 Schwitzer G. A review of features in Internet consumer health decision-support tools. J Med Internet Res. 2002 Apr-Nov;4(2):E11.

7 Fox S, Fallows D. Health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access. Pew Internet & American Life Project. 2003 16 Jul.

8 IOM (Institute of Medicine). Health literacy: a prescription to end confusion. Washington, DC: National Academy Press. 2004.

9 Zielstorff RD. Controlled vocabularies for consumer health. Biomed Inform. 2003 Aug-Oct;36(4-5):326-33.

10 Zeng Q, Kogan S, Ash N, Greenes RA, Boxwala AA. Characteristics of consumer terminology for health in-formation retrieval. Meth Inf Med. 2002;41(4):289-98.

11 Smith CA, Stavri PA, Chapman WW. In their own words? A terminological analysis of e-mail to a cancer information service. Proc AMIA Symp. 2002:697-701.

12 Tse T, Soergel D. Exploring medical expressions used by consumers and the media: an emerging view of consumer health vocabularies. Proc AMIA Symp. 2003: 564-8.

13 Young CG, Barger JD. Learning medical terminology step by step. Saint Louis: C.V. Mosby. 1975.

14 Carroll JM. What's in a name; an essay in the psychology of reference. New York: W.H. Freeman. 1985.

15 Furnas GW, Landauer TK, Gomez LM, Dumais ST. The vocabulary problem in human-system communication. Communications of the ACM. 1987;30(11):964-71.

16 Osborne H. Health Literacy From A To Z: Practical Ways To Communicate Your Health: Jones & Bartlett Pub; 2004.

17 Spiro D, Heidrich F. Lay understanding of medical terminology. J Fam Prac. 1983;17(2):277-9.

18 Gibbs RD, Gibbs PH, Henrich J. Patient understanding of commonly used medical vocabulary. J Fam Prac. 1987;25(2):176-8.

19 Sugarman J, Butters RR. Understanding the patient: medical words the doctor may not know. NCMJ. 1985 Jul:415-7.

20 Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med. 1995 Oct;10(10):537-41.

21 Miller N, Lacroix EM, Backus JE. MedlinePlus: building and maintaining the National Library of Medicine's consumer health Web service. Bull Med Libr Assoc 2000;88(1):11-7.

22 Divita G, Tse T, Roth L. Failure analysis of MetaMap Transfer (MMTx). Medinfo;2004:763-7.

23 Ogden J, Branson R, Bryett A, Campbell A, Febles A, Ferguson I, Lavender H, Mizan J, Simpson R, Tayler M. What's in a name? An experimental study of patients' views of the impact and function of a diagnosis. Fam Pract. 2003 Jun;20(3):248-53.