Viewpoint Paper ■

Exploring and Developing Consumer Health Vocabularies

QING T. ZENG, PHD, TONY TSE, PHD

A b s t r a c t   Laypersons ("consumers") often have difficulty finding, understanding, and acting on health information due to gaps in their domain knowledge. Ideally, consumer health vocabularies (CHVs) would reflect the different ways consumers express and think about health topics, helping to bridge this vocabulary gap. However, despite the recent research on mismatches between consumer and professional language (e.g., lexical, semantic, and explanatory), there have been few systematic efforts to develop and evaluate CHVs. This paper presents the point of view that CHV development is practical and necessary for extending research on informatics-based tools to facilitate consumer health information seeking, retrieval, and understanding. In support of the view, we briefly describe a distributed, bottom-up approach for (1) exploring the relationship between common consumer health expressions and professional concepts and (2) developing an open-access, preliminary (draft) "first-generation" CHV. While recognizing the limitations of the approach (e.g., not addressing psychosocial and cultural factors), we suggest that such exploratory research and development will yield insights into the nature of consumer health expressions and assist developers in creating tools and applications to support consumer health information seeking.

■ **J Am Med Inform Assoc.** 2006;13:24–29. DOI 10.1197/jamia.M1761.

A Case for Consumer Health Vocabularies

Although it has been long recognized that laypersons ("consumers") and health care professionals express and think about health-related concepts differently, such mismatches in language continue to hinder effective communication, health information seeking, and, ultimately, informed decision making. This vocabulary gap is becoming a more serious problem as consumers increasingly explore health-related information resources on their own and assume greater responsibility for personal health care. Further, technical terminology (or jargon) is a barrier to health literacy, defined as "the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions."[1] Low health literacy has been associated with poor health outcomes (see ref. 2 for a literature review).

Affiliations of the authors: Department of Radiology, Decision Systems Group, Brigham and Women's Hospital, Harvard Medical School, Boston, MA (QTZ); Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Department of Health and Human Services, Bethesda, MD (TT).

Correspondence and reprints: Qing T. Zeng, PhD, Department of Radiology, Decision Systems Group, Thorn 309, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115; e-mail: <qzeng@dsg.bwh.harvard.edu>.

In the general biomedical literature, research on consumer understanding of medical words and concepts has focused primarily on lists of discrete terms in various specialties, e.g., 17 oncology terms,[3] six word pairs among emergency department patients,[4] and 50 "common" medical terms among surgical patients and family members.[5] While such studies consistently confirm that consumers have difficulty comprehending medical jargon, they do not provide a methodology for systematically identifying alternative words or phrases (collectively, "expressions") that consumers are likely to use, recognize, and understand.

Concerted informatics research on "consumer health vocabularies" (CHVs) began relatively recently.[6–9] In 2001, a distinguished panel at MEDINFO enumerated several significant challenges in creating a CHV and concluded:

> The development of a consumer vocabulary should be based on research that includes consumer information needs and consumers' ways of talking about expressing those needs.... The healthcare consumer then would not need to know the language of the healthcare provider in order to effectively find relevant information resources, and he/she could search for resources in his/her own language....[10]

However, as Lewis and colleagues[10] also point out, "to presume that [the language of consumers] contains a knowable stable vocabulary and grammar similar in structure to that of the formal languages imposes a professional structure on a very personal experience." Indeed, consumers are likely to use a combination of "everyday language," technical terms (with or without knowledge of the underlying concepts), and various explanatory models, all influenced by psychosocial and cultural variations, in discourse about health topics. For example, Kleinman and colleagues[11] described a case in which an older woman with pulmonary edema was inducing vomiting and urinating frequently in bed while hospitalized. When asked about her apparently "unusual" behavior, she replied that because the medical team told her that she had

"water in the lungs," she was trying to help remove it. In this case, the descriptive phrase used by well-meaning professionals elicited an explanatory model based on a plumbing metaphor. Interestingly, her father and husband were plumbers by profession. This illustrates that the language and mental models used by consumers represent a complex phenomenon at multiple levels (e.g., lexical, semantic, and explanatory) that is influenced by countless internal and external factors, varying from person to person and situation to situation.

Nevertheless, an understanding of some "common" health expressions, concepts, or explanatory models shared within discourse groups (e.g., cultural, geographical, and familial) will help researchers build a foundation upon which "personal health vocabularies" might ultimately be developed to reflect personal, individualized health-related constructs. For purposes of the present discussion, CHV refers to a collection of expressions, concepts, attitudes, and beliefs observed to be used by most members of a consumer discourse group to communicate about health-related issues (e.g., online users of a consumer health Web site, members of a discussion forum). While researchers and developers are likely to impose some organization on CHVs for specific, pragmatic tasks (e.g., developing a bridge or crosswalk between consumer expressions and professional terms to aid information retrieval), it is recognized that such structures are artificial and, at best, rough approximations of the underlying "reality." The case recounted above is a reminder that continued research on CHVs must address not only the potential benefits, but also potential risks, such as increasing confusion or seemingly validating incorrect explanatory models.

While subsets of terms related to CHVs have been identified and characterized in recent exploratory research[12–14] and the effects of the vocabulary gap on health information retrieval have been documented,[15] there has been little systematic research and development on CHVs (see ref. 16 for a comprehensive review). As a result, no comprehensive, nonproprietary CHV is currently available to researchers and developers, and no evaluation or validation of CHVs has been reported.

This paper presents the point of view that, despite the substantial theoretical and practical challenges in defining and developing a "standardized consumer vocabulary,"[10] practical methods are available for leveraging existing resources to begin exploring and developing a "first-generation" open-access CHV, an important, if not essential, tool for consumer health informatics. Based on preliminary research, the authors contend that:

- Expressions used by consumers to describe health-related concepts and relationships among such concepts frequently differ from those of professionals;
- Differences also exist at the conceptual and explanatory levels;
- Such consumer-oriented, health-related forms and concepts are sufficiently consistent to be identified and characterized as belonging to distinct CHVs, at least among discourse groups for particular tasks;
- Consumer health informatics research and applications development will benefit from the development of CHVs.

Applied research on ways to develop and update CHVs systematically will yield insights into the nature of the vocabulary gap as well as practical solutions for narrowing that gap. The purpose of this paper is to increase awareness of the need for such a vocabulary, stimulate dialogue and action on the development of open-access CHVs, and ultimately, help meet the challenges of bridging the gap between consumers and relevant, appropriate health-related information.

Consumer and Professional Language in the Health Domain

Consumers use words and phrases (expressions) to describe health-related concepts that frequently differ from those used by professionals. While health care specialists share foundational domain knowledge based on formal education and professional experience, laypersons have some socially and culturally derived notions of health and illness[17] acquired from formal and informal sources (e.g., media exposure) and unique personal experiences. Thus, whereas terminology facilitates discourse among professionals in their everyday routines, it can be incomprehensible or confusing to consumers.

During face-to-face communication between consumers and professionals, the interactive nature of such encounters allows for the possibility of negotiating ideas at a "common level of discourse." For example, either person may ask for clarification or repeat important points when confusion is detected. However, "static" public-oriented media such as brochures, pamphlets, newspapers, and television, where many people obtain their health information, rely on a "one-size-fits-all" approach based on the knowledge and judgment of individual authors. Finally, it is even a challenge to customize information dynamically to facilitate comprehension for online users.

Further, during disintermediated interactions, such as via computer or printed form, consumers tend to "fill in" gaps in comprehension (correctly or incorrectly), using their own knowledge, experience, and preferences, as lay language shares some terms and concepts with professional language. In the best case, consumers will infer the meaning of a term correctly and even use the term (e.g., *attention deficit disorder*) as their own in future health communication. Among the worst case scenarios, consumers will misinterpret or (mis)associate a term with context or connotations not intended by the content provider or author. Alternatively, consumers may recognize or use a technical form, but associate it with a concept in lay usage rather than one from a professional health care domain (e.g., *depression* for the emotion *sadness*, rather than a psychological condition such as *depressive disorder*). Consumers who do not or cannot fill in the gaps at all (e.g., due to low health literacy) would simply be confused, the information having been "lost in translation."

When producing words to describe health-related concepts, a lay person may use terms such as *hair loss* and *heart failure* without knowing their technical definitions or use general language expressions to describe familiar concepts (e.g., *loss of appetite* for *anorexia* and *pain killer* for *analgesic*). The range of lay expressions seems to vary from general and descriptive (e.g., *device to look inside my ear* for *otoscope*) to specific, but colloquial (e.g., *sugar* for *diabetes*). Thus, lay discourse on the health-related topics often includes a combination of technical terminology and general language expressions, with many

possible interpretations based on individual, contextual, societal, and cultural associations. The challenge is to sort out the different ways consumers communicate within distinct discourse groups and map the common, shared expressions and contexts to the more constrained, specialized language of professionals, when appropriate.

Recent informatics studies on CHVs have contributed insights about the differences between the languages and vocabularies of consumers and professionals. While these studies have shared general approaches and findings, such as use of consumer-authored text, automated mapping to controlled medical vocabularies, and analyzing text that could not be mapped to controlled vocabularies, many of the details vary. For example, sources of consumer-authored texts included e-mail messages,[12,14] online health discussion group postings,[13] queries submitted to consumer health Web sites,[6,8,9,15,18,19] and folk terms from the Dictionary of American Regional English project.[8] Two studies focused on particular medical domain areas, cancer[12] and diabetes,[8] while others reviewed expressions from the general health domain. With regard to identifying and mapping expressions to professional concepts, although one study used consumer surrogates to identify consumer health expressions manually,[13] all the studies mapped consumer expressions to the components of the National Library of Medicine (NLM) Unified Medical Language System (UMLS) Metathesaurus. (Note that Brennan and Aronson[14] focused on nursing vocabularies while Patrick and colleagues[8] also used the Eurodicautom http://europa.eu.int/eurodicautom/Controller and the European Commission Multilingual Glossary of Popular and Technical Medical Terms in Nine European Languages http://allserv.rug.ac.be/~rvdstich/eugloss/welcome.html.)

Taken together, these studies examined large numbers of consumer utterances (i.e., hundreds of thousands of tokens) and consistently found that between 20% and 50% of consumer health expressions were not represented by professional health vocabularies. A subset of these unrepresented expressions (i.e., hundreds to thousands) underwent human review.[13,15,18] The majority of unmapped consumer health expressions represented either (1) forms for existing concepts in health vocabularies not found algorithmically (e.g., due to typos, descriptive phrases, or meronymy) or (2) expressions for legitimate health concepts not represented in the health vocabularies. Because each of these studies used utterances from artifactual sources (text), the intended meaning of the expressions had to be inferred by the researcher, using contextual cues. Follow-up research is needed to explore the actual meaning(s) attributed to the health expressions by consumers from different discourse groups and their explanatory models.

Several studies that have used qualitative methods, such as direct observation, semistructured interview, and semantic analysis of patient-clinician dialogue, suggest that consumer-professional language mismatches occur at the conceptual, relational, and mental model levels.[15,20,21] These findings are consistent with earlier work that observed conceptual differences,[22] cultural influences,[11] and differences in knowledge representation (illness explanatory model used by consumers and the disease model used by professionals[23]).

Stability of Lay Health Language

The utility of computable vocabularies depends on the stability and organization of the languages they represent. Typically, vocabularies consist of three high-level components: (1) surface forms, such as alphanumeric strings; (2) underlying meanings representing distinct concepts; and (3) relationships between them (i.e., form-form, form-concept, or concept-concept). In an ideal vocabulary, these components would be encoded in a completely logical, principled, and controlled manner. However, as discussed previously, consumer health language is as flexible, dynamic, and complex as general language. Thus, the feasibility of developing useful CHVs relies on the fundamental question about the *stability* of lay health language. That is, are the health-related forms, concepts, and relationships familiar to and used by consumers sufficiently consistent to allow for the development of a useful vocabulary? The available evidence, reviewed below, suggests that aspects of lay language in the health domain do have a sufficient level of consistency (at least among typical online active consumers) to support CHVs.

First, researchers have identified and applied generalized patterns of concepts to characterize health terms and expressions for consumers. For instance, "reverse medical dictionaries" list common words and phrases associated with their professional counterparts, as finding aids for consumers.[24] Scott and Weiner[25] identified several types of professional terms that are difficult for lay comprehension:

- Difficult general language words having the same meanings as technical terms (e.g., *alleviate*, *apprehensive*, and *intermittent*);
- Technical terms requiring domain knowledge to understand (e.g., *angina*, *bronchus*, and *diastolic*); and
- General language words with different technical meanings (e.g., *negative*, "not constructive" vs. "absence of a condition"; *sob*, "weeping" vs. "shortness of breath," and *anorexia*, "eating disorder" vs. "loss of appetite").

Other researchers have explored the effects of using consumer or professional language during medical consultations. For example, while patients preferred to receive their diagnoses in technical terms rather than as lay expressions,[26] they were more satisfied with encounters in which a physician adjusted his/her language to match theirs, including slang, than when the physician used technical jargon.[27]

Second, the authors' own research suggests that common levels of discourse, in which a majority of people expresses similar concepts using particular "consensus" forms, can be found among narrowly defined populations for specific tasks. The analysis of over 12 million queries submitted to the NLM consumer health information system, MedlinePlus http://www.medlineplus.gov from October 2002 to September 2003, indicated that a small number of words and phrases occurred very frequently while a majority of expressions appeared only once or twice. For instance, of all the occurrences of expressions likely to map to UMLS concept C0022650 (*kidney calculi*) in the data set, *kidney stones* accounted for over 95% of them. Further, *wrist* and *carpal tunnel syndrome* co-occurred quite often, suggesting a short semantic distance between the two in consumer mental models (e.g., *carpal tunnel syndrome* occurs_in *wrist*).

Third, general language dictionaries provide standard definitions of many common lay health expressions. For example, *overweight* is defined by the *American Heritage Dictionary of the English Language, Fourth Edition*[28] as "more weight than is normal, necessary, or allowed." When referring to health, this "generic" lay notion of *overweight* may be mapped to two professional concepts of greater specificity in the UMLS: (1) *Obesity* and (2) *Overweight (BMI <30)*, where *obese* is defined as body mass index (BMI) ≥30 and *overweight* as BMI between 25 and 30. Thus, the conceptualization of *overweight* described in the dictionary entry is less specific than definitions used by professionals, suggesting that dictionary entries (i.e., definitions and senses), to the extent that they represent shared meanings within a larger discourse group, are sufficiently stable as a starting point for CHVs. Clearly, "standard" dictionary definitions alone are unsuitable for CHV development; since meaning is constructed by individuals, it is important for CHVs to reflect social, cultural, and contextual variability.

While the evidence that a core set of health-related expressions is used consistently by consumers has been highlighted, the authors acknowledge the inherent expressive variability in lay health language. Unlike health care professionals who are exposed to standardized terminology through formalized training and experience, consumers come from a variety of cultural, social, and educational backgrounds and have different family or personal health histories and experiences. It would be unrealistic to expect otherwise. Of course, variability also exists in professional health language (e.g., synonyms, ambiguity, and vagueness). For instance, the distinction between *diagnosis* and *finding* is not always clear.[29] Thus, while variability in lay health language increases the complexity of CHV development, it should not be viewed as prohibitive.

Uses of Consumer Health Vocabularies

Eysenbach[30] observed that "consumer health informatics is designed to empower consumers by putting health information into their hands, including information on their own health, such as diagnoses, lab results, personal risk factors, and prescribed drugs." For example, applications allow patients to enter their data directly into electronic medical records and to send e-mail to their doctors.[31,32] An implication is that computer systems now need to process health-related data directly entered by consumers, such as queries or self-reported symptoms. Conversely, general or personal health information presented to consumers should be comprehensible. To achieve such two-way communication effectively requires CHV support. Three actual cases illustrate the bridging role of CHVs between consumers and health applications or health information for different types of tasks:

- **Information retrieval.** While consumer-formulated query terms often do not map to the UMLS, pilot studies show that professional terms are likely to produce better search results than consumer expressions.[8,33,34] CHVs would facilitate automated mapping of consumer-entered queries to technical terms (e.g., *belly button* → *navel*; *weight loss surgery* → *bariatric surgery*).
- **Medical records.** While medical records and test results are becoming available to patients, they frequently contain jargon. A CHV could supplement these terms with consumer-

friendly display (CFD) names to help patients understand the terms (e.g., *malignant neoplasm* → *cancer*; *pruritus* → *itching*).
- **Health Care Applications.** Patients may enter consumer expressions such as *sun poisoning* or *hole in the heart* when using computer-assisted interview,[31,32] decision support,[35] and disease management[36] applications. CHVs would facilitate automated mapping of these expressions to concepts and enable subsequent analysis and response.

Development of Consumer Health Vocabularies

Even though lay expressions have been researched and developed in other domains (e.g., health communication), the resulting free-text lists and dictionaries are not sufficient for consumer informatics applications, any more than medical dictionaries could replace controlled medical vocabularies in clinical informatics systems. A computerized collection of health expressions derived from actual consumer utterances (authored by consumers), linked to professional concepts, and reviewed and validated by professionals and consumers, would be a first step toward providing greatly needed interventions for assisting consumers during interactions with health information.

Whereas a general notion of CHVs was described in the introduction, an operational definition of a "first-generation" CHV is a collection of forms used in health-oriented communication for a particular task or need (e.g., information retrieval) by a substantial percentage of consumers from a specific discourse group and the relationship of the forms to professional concepts. Note that "consumer concepts" and "lay mental/explanatory models" will be incorporated in future "n-generation" CHVs. Such a pragmatic definition embodies the core features of a CHV, while recognizing longer term issues (e.g., generalizability of any CHV to other discourse groups and tasks). That is, preliminary CHV development may begin immediately, using existing informatics technologies, while research on broader issues of consumer health lexicology/terminology and various sociocultural influences is conducted in parallel.

In a pilot study to develop a first-generation CHV, the authors experimented with such a phased, distributed, user source–based approach.[19] The initial phase focused on identifying consumer forms for "standard" health-related concepts or CFD display names. Consumer concepts and relationships not represented in existing health vocabularies will be handled in subsequent phases. This sequence reflects the degree of complexity required to develop each of the three vocabulary components, as well as the conventional steps of vocabulary development.[37] Text strings (surface forms) are generally less difficult to observe and more tractable to computational processing than conceptualizations and semantic relations in mental models.

A two-step procedure was used to assign CFD names to the most frequently occurring concepts observed among MedlinePlus queries. In the first step, commonly used consumer expressions were mapped to the UMLS Metathesaurus (2004AA version), using lexical processes. During the second step, two reviewers (with a third as tiebreaker) assessed expressions matched to each of the most common UMLS concepts and identified candidate CFD names. Groups of reviewers discussed and selected CFD names. Over a six-month

period, up to seven reviewers from NLM and Brigham and Women's Hospital (including the authors), with diverse backgrounds in consumer health, computer science, and terminology development, participated. Researchers from several other organizations were consulted on difficult cases. Initially a set of 425 concepts was reviewed and a preliminary evaluation of 34 concepts showed that CFD names led to better comprehension.[19] As of June 2005, 1,000 concepts have been reviewed; we are also exploring other approaches such as the one reported by Fung et al.[38] to validate the CFD names.

The authors also presented the preliminary results of the distributed CHV development experiment at the Consumer Health Informatics Working Group meetings of the American Medical Informatics Association (AMIA) and International Medical Informatics Association (IMIA) at the 2004 MEDINFO conference in San Francisco, which led to discussions with several interested groups.

The pilot study tested the feasibility of the phased, distributed, user source–based approach for CHV development. Consumer health research spans specialties and domains. Thus, CHV development requires not only the involvement of physicians, nurses, and informaticians, "but linguists, medical librarians, and representatives of other health disciplines and, of course, patients themselves."[16]

Because health care is a vital issue for public health, CHVs should be open access to encourage widespread use and further development and refinement. Researchers and developers would be encouraged to contribute innovations and new data derived from alternative or local uses in the community back to the CHV. There are many examples of successful open-access projects in computing, and the concept has begun to gain support in the biomedical informatics domain.[39] For example, the availability of the UMLS to the informatics community has stimulated and facilitated numerous research and development projects. One difference is that, while the UMLS necessarily relies on existing vocabularies, many of which are proprietary, we have an opportunity to develop fully open-access CHVs, as there are few legacy issues.

Last, as advocated by the MEDINFO 2001 panel,[10] CHV development should be based on actual consumer expressions. Mining sources of authentic consumer utterances (in aggregate) over time, while preserving individual privacy, will help account for variability in general language and "drift" in the forms and concepts used across discourse groups. Since CHVs will surely be refined, enriched, and specialized over time, this approach also has the advantage of being generalizable. Direct observation and interviews may be used to complement or validate the text mining approach to the study of forms and mapping of forms to concepts, although they are not appropriate for investigating large numbers of terms or concepts.

Conclusion

Patel and Kaufman[40] have suggested that medical informatics is a "local science of design." That is, the field is "local" to specific aspects or parts of the broader domain (contrast with "global") and "design" because it looks toward pragmatic outcomes: "in practice, design is strongly bound by domain-specific constraints and grounded in contexts in which an artifact is to be used." The view of developing a first-

generation CHV for particular consumers with specific needs is consistent with this thesis. Development of a CHV should not be hampered by the complexities of the language used by consumers, but should embrace the challenge of making health-related information more accessible to members of the public and better suited to their needs and individual circumstances. "One of the goals of a local science of design is to discover what works and then determine why some things work and others don't. A working system is an outcome not merely of technology but of the social and cognitive processes of integrating such a system into daily workflow."[39]

A number of studies suggest that the vocabulary gap between consumers and professionals is a substantial barrier to health information access for laypersons. A first-generation CHV is likely to serve as a platform for incremental improvements in consumer health information seeking while providing new insights into the overall problem of bridging lay forms, concepts, and relations with the medical domain. Simultaneously, research on the contextual, sociocultural, and other factors that affect how laypersons express and think about health topics will illuminate the other dimensions of CHVs. By tackling small pieces of the larger problem, through a concerted, interdisciplinary, and open-access effort by the consumer health informatics research community, it is anticipated that a convergence in understanding of the overall phenomenon will result.

References ∎

1. Ratzan SC, Parker RM. Introduction. In: National Library of Medicine current bibliographies in medicine: health literacy, Selden CR, Zorn M, Ratzan SC, Parker RM, editors. NLM Pub No CMB 200-1. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services, 2000.
2. Institute of Medicine. Health literacy: a prescription to end confusion. Washington DC: National Academy Press; 2004. pp. 81–103.
3. Chapman K, Abraham C, Jenkins V, Fallowfield L. Lay understanding of terms used in cancer consultations. Psychooncology. 2003;12:557–66.
4. Lerner EB, Jehle DV, Janicke DM, Moscati RM. Medical communication: do our patients understand? Am J Emerg Med. 2000;18: 764–6.
5. Spees CM. Knowledge of medical terminology among clients and families. Image J Nurs Sch. 1991;23:225–9.
6. McCray AT, Loane RF, Browne AC, Bangalore AK. Terminology issues in user access to Web-based medical information. Proc AMIA Symp. 1999;107–11.
7. Marshall PD. Bridging the terminology gap between health care professionals and patients with the Consumer Health Terminology (CHT) Thesaurus. Proc AMIA Symp. 2000;1082.
8. Patrick TB, Monga HK, Sievert MC, Hall JH, Longo DR. Evaluation of controlled vocabulary resources for development of a consumer entry vocabulary for diabetes. J Med Internet Res. 2001;3:e24. Retrieved 2005 Jun 21. Available from: http://www.jmir.org/2001/3/e24/.
9. Zeng Q, Kogan S, Ash N, Greenes RA. Patient and clinician vocabulary: how different are they? Medinfo. 2001;10:399–403.
10. Lewis D, Brennan PF, McCray AT, Tuttle M, Bachman J. If we build it, they will come: standardized consumer vocabularies. Medinfo. 2001;10:1530.
11. Kleinman A, Eisenberg L, Good B. Culture, illness, and care: clinical lessons from anthropologic and cross-cultural research. Ann Intern Med. 1978;88:251–8.
12. Smith CA, Stavri PZ, Chapman WW. In their own words? A terminological analysis of e-mail to a cancer information service. Proc AMIA Symp. 2002;697–701.

13. Tse T, Soergel D. Exploring medical expressions used by consumers and the media: an emerging view of consumer health vocabularies. Proc AMIA Symp. 2003;674–8.

14. Brennan PF, Aronson AR. Towards linking patients and clinical information: detecting UMLS concepts in e-mail. J Biomed Inform. 2003;36:334–41.

15. Zeng Q, Kogan S, Ash N, Greenes RA, Boxwala AA. Characteristics of consumer terminology for health information retrieval. Methods Inf Med. 2002;41:289–98.

16. Zielstorff RD. Controlled vocabularies for consumer health. J Biomed Inform. 2003;36:326–33.

17. Wood ML. Naming the illness: the power of words. Fam Med. 1991;23:534–8.

18. McCray AT, Tse T. Understanding search failures in consumer health information systems. Proc AMIA Symp. 2003;430–4.

19. Zeng QT, Tse T, Crowell J, Divita G, Roth L, Browne AC. Identifying consumer-friendly display (CFD) names for health concepts. Proc AMIA Symp. 2005;869–73.

20. Kogan S, Zeng Q, Ash N, Greenes RA. Problems and challenges in patient information retrieval: a descriptive study. Proc AMIA Symp. 2001;329–33.

21. Slaughter L, Soergel D. How physicians' answers relate to health consumers' questions. Proc AMIA Symp. 2003;28–39.

22. Ogden J, Boden J, Caird R, et al. You're depressed no I'm not: GPs' and patients' different models of depression. Br J Gen Pract. 1999;49:123–4.

23. Patel VL, Arocha JF, Kushniruk AW. Patients' and physicians' understanding of health and biomedical concepts: relationship to the design of EMR systems. J Biomed Inform. 2002;35:8–16.

24. Hamilton B, Guidos B. The medical word finder: a reverse medical dictionary. New York: Neal Schuman Publishers; 1987.

25. Scott N, Weiner MF. 'Patientspeak': an exercise in communication. J Med Educ. 1984;59:890–3.

26. Ogden J, Branson R, Bryett A, Campbell A, Febles A, Ferguson I, et al. What's in a name? An experimental study of patients' views of the impact and function of a diagnosis. Fam Pract. 2003;20:248–53.

27. Williams N, Ogden J. The impact of matching the patient's vocabulary: a randomized control trial. Fam Pract. 2004;21:630–5.

28. The American Heritage Dictionary of the English Language. 4th ed. New York: Houghton Mifflin; 2000.

29. Greenes RA, McClure RC, Pattison-Gordon E, Sato L. The findings–diagnosis continuum: implications for image descriptions and clinical databases. Proc Annu Symp Comput Appl Med Care. 1992;383–7.

30. Eysenbach G. Consumer health informatics. BMJ. 2000;320: 1713–6.

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

32. Wald JS, Rind D, Safran C, Kowaloff H, Barker R, Slack WV. Patient entries in the electronic medical record: an interactive interview used in primary care. Proc Annu Symp Comput Appl Med Care. 1995;147–51.

33. Sievert ME, Patrick TB, Reid JC. Need a bloody nose be a nosebleed? or, lexical variants cause surprising results. Bull Med Libr Assoc. 2001;89:68–71.

34. Plovnick RM, Zeng QT. Reformulation of consumer health queries with professional terminology: a pilot study. J Med Internet Res. 2004;6:e27.

35. Bouhaddou O, Lambert JG, Miller S. Consumer health informatics: knowledge engineering and evaluation studies of medical HouseCall. Proc AMIA Symp. 1998;612–6.

36. Ralston JD, Revere D, Robins LS, Goldberg HI. Patients' experience with a diabetes support programme based on an interactive electronic medical record: qualitative study. BMJ. 2004; 328:1159.

37. Soergel D. Indexing languages and thesauri: construction and maintenance. Los Angeles: Melville; 1974.

38. Fung KW, Hole WT, Nelson SJ, Srinivasan S, Powell T, Roth L. Integrating SNOMED CT into the UMLS: an exploration of different views of synonymy and quality of editing. J Am Med Inform Assoc. 2005;12:486–94.

39. McDonald CJ, Schadow G, Barnes M, Dexter P, Overhage JM, Mamlin B, McCoy JM. Open Source software in medical informatics—why, how and what. Int J Med Inf. 2003;69:175–84.

40. Patel VL, Kaufman DR. Science and practice: a case for medical informatics as a local science of design. J Am Med Inform Assoc. 1998;5:489–92.