

Search terms and a validated brief search filter to retrieve publications on health-related values in Medline: a word frequency analysis study

Mila Petrova,^{1,2} Paul Sutcliffe,² K W M (Bill) Fulford,³ Jeremy Dale²

► Additional tables and information are published online only. To view these files please visit the journal online (www.jamia.bmj.com/content/19/3.toc).

¹Egenis (ESRC Centre for Genomics in Society), University of Exeter, Exeter, UK

²Health Sciences Research Institute, Warwick Medical School, University of Warwick, Coventry, UK

³Institute of Clinical Education, Warwick Medical School, University of Warwick, Coventry, UK

Correspondence to

Mila Petrova, Egenis, ESRC Centre for Genomics in Society, Byrne House, St German's Road, Exeter EX4 4PJ, UK; mp320@exeter.ac.uk

Project carried out at Warwick Medical School.

Received 12 March 2011
Accepted 11 July 2011
Published Online First
16 August 2011

ABSTRACT

Objective Healthcare debates and policy developments are increasingly concerned with a broad range of values-related areas. These include not only ethical, moral, religious, and other types of values 'proper', but also beliefs, preferences, experiences, choices, satisfaction, quality of life, etc. Research on such issues may be difficult to retrieve. This study used word frequency analysis to generate a broad pool of search terms and a brief filter to facilitate relevant searches in bibliographic databases.

Methods Word frequency analysis for 'values terms' was performed on citations on diabetes, obesity, dementia, and schizophrenia (Medline; 2004–2006; 4440 citations; 1 110 291 words). Concordance® and SPSS 14.0 were used. Text words and MeSH terms of high frequency and precision were compiled into a search filter. It was validated on datasets of citations on dentistry and food hypersensitivity.

Results 144 unique text words and 124 unique MeSH terms of moderate and high frequency (≥ 20) and very high precision ($\geq 90\%$) were identified. Of these, 19 text words and seven MeSH terms were compiled into a 'brief values filter'. In the derivation dataset, it had a sensitivity of 76.8% and precision of 86.8%. In the validation datasets, its sensitivity and precision were, respectively, 70.1% and 63.6% (food hypersensitivity) and 47.1% and 82.6% (dentistry).

Conclusions This study provided a varied pool of search terms and a simple and highly effective tool for retrieving publications on health-related values. Further work is required to facilitate access to such research and enhance its chances of being translated into practice, policy, and service improvements.

BACKGROUND AND SIGNIFICANCE

Healthcare debates and policy developments are increasingly concerned with a broad range of values-related areas. Most prominently, these include medical ethics and patients' rights. But values-related issues in healthcare extend much further, to issues such as: recognition of patients' wishes and preferences in shared decision making^{1–3}; respect for difference and diversity in relation to ethnicity, culture, religious beliefs, sexual orientation, and lifestyle^{4–5}; issues of personalized healthcare,^{6–7} self-management,^{8–9} patient choice,^{10–11} and satisfaction^{12–15}; the use of patient-reported outcomes in evaluations of treatments and estimates of cost-effectiveness^{14–15}; community engagement^{16–18}; differences in perspective and priorities in multidisciplinary teams^{2–19}; and many more. Therefore there is an

increasing need for identification of relevant research. Yet achieving this is not straightforward.

Searches based on the word 'value' tend to have low precision and/or low sensitivity (see box 1: glossary). If the text word 'values' is used in Medline, technical or statistical phrases tend to be identified: 'abnormal laboratory values', 'the values of systolic and diastolic blood pressure', 'values of <60 ml/min per 1.73 m²', 'evidence-based target values'. In contrast, relevant thesaurus terms (see box 1) tend to have 100% precision but very low sensitivity. For example, Social Values [MeSH] AND Diabetes Mellitus [MeSH] retrieved 23 publications for all preceding years out of 260 337 publications indexed with Diabetes Mellitus [MeSH] (search run Mar 2011; Medline; PubMed interface). Such findings indicate the need for longer search strategies, incorporating search terms for specific values issues.

Health-related values

An immediate problem for developing search strategies for publications on values is the variety of meanings of the word 'value'. The *Oxford English Dictionary* lists seven main meanings for it as a noun, a further eight as a verb, and a total of 40 sub-meanings excluding examples, compounds, and recent additions.²² The philosopher G H von Wright described several hundred uses of the term in his seminal *The Varieties of Goodness*.²³ Here, we use 'values' to include anything positively or negatively weighted as a guide to decisions and action. This is consistent with uses in analytic philosophy,²⁴ decision analysis,²⁵ and health economics.²⁶ With values construed in this way, health-related values are explored from a variety of disciplinary perspectives, including medical ethics, the sociology of health and medicine, health economics, health psychology, medical anthropology, and the medical humanities.

There are no established definitions and classifications of health-related values to use as a starting point in developing relevant search strategies. Box 2 outlines some of the ways in which we operationalized the concept through instructions and taxonomies. These were used to guide the selection of keywords for searches to compile the derivation datasets, make inclusion/exclusion decisions, anchor judgments of saturation, and resolve inter-rater disputes. Online supplementary file 1, section 1 presents the background work underpinning these operationalizations. Briefly, it included researching theoretical literature on values, exploring our gradually expanding collection of

Box 1 Glossary of terms used in text and supplementary files (definitions draw primarily on PubMed Help, the website of the Concordance® software, the *Oxford English Dictionary*, Wong *et al*²⁰ and Jenkins²¹)

Concordance—an alphabetical arrangement of the principal words contained in a text, with citations of the passages in which they occur (OED). In this paper, we use it to refer to alphabetically arranged lists of words and their frequency counts. **Concordance**® is a software product that generates such word lists and frequency counts. A ‘full concordance’ is one that includes all words appearing in a textual corpus. If required, words can be excluded from a full concordance by means of a **stop list**. A ‘selective concordance’ includes only the frequencies for words from a **pick list** (a user-defined list of words).

Text words—search terms that are not sourced from an existing thesaurus. They can be used more flexibly than thesaurus terms, by applying truncations and wildcards, performing proximity searches, etc, but are likely to produce a large number of false positives. Also called ‘natural language terms’.

Frequency—in the context of this paper, represents the number of times a text word appears in the analyzed dataset and is used as proxy for sensitivity. It may overestimate sensitivity, as it reflects the number of occurrences of a word rather than the number of citations in which it is found.

Keywords—in the context of this paper, used generically to include text words, MeSH terms, and words reflecting aspects of the concept of value whether or not they are used as search terms. Some databases use ‘keywords’ to denote thesaurus terms (see below).

MeSH (medical subject headings)—Medical subject headings comprise the controlled vocabulary of biomedical terms of the US National Library of Medicine, publisher of Medline. A journal article in Medline is typically assigned 10–12 MeSH, as specific as possible. In this paper (as is the convention more generally), we have distinguished MeSH terms from text words through the use of capital letters.

‘Open’ searches (non-standard term—convention used in this paper)—searches in which only health topic search terms were used, with no search terms for values contents added.

Precision—the number of relevant articles retrieved as a proportion of the total number of articles retrieved by a particular search (equivalent to ‘positive predictive value’ in diagnostic test terminology).

Retrieval—the set of references returned by a search. Also called ‘yield’.

Search filter—a predefined search strategy traditionally designed to retrieve high-quality evidence (eg, from randomized controlled trials, systematic reviews) and/or evidence pertaining to one of the generic areas of clinical concern (diagnosis, prognosis, etc). Search filters are a type of search strategy—one that has been developed through a rigorous process and makes claims to facilitating optimal retrieval, along predefined dimensions (eg, sensitivity, precision, etc).

Sensitivity—the number of relevant articles identified by a search as a proportion of the total number of relevant articles. Also referred to as ‘recall’.

Specificity—the number of false positives avoided by a search as a proportion of the total number of false positives.

Subheadings—MeSH subheadings are used with MeSH terms to further qualify a topic. For example, an article discussing legal aspects of Personal Autonomy will be indexed as personal autonomy/legislation and jurisprudence.

Thesaurus terms—preselected terms that are used to index publications in databases, catalogs, registers, etc and which enable better targeted and precise searches. Also called ‘controlled vocabulary terms’, ‘index(ing) terms’, ‘descriptors’, etc.

empirical studies on health-related values, gathering ideas from colleagues and participants in values awareness training sessions, and accessing the wider social discussion (eg, through media materials and artwork).

Search strategies already available

We identified a number of search strategies designed for other purposes that could support searches for publications on health-related values. Kahn and Ninomiya have produced an extensive guide on searching for bioethics and medical ethics literature.²⁷ Progress has been made in the development and evaluation of search filters for qualitative research.^{20 28 29} Systematic reviews on psychological, psychosocial, and patient involvement issues are also a source of relevant search strategies (see, eg, reviews of The Cochrane Consumers and Communication Group³⁰). Closest to a comprehensive values search strategy as envisaged here were McNally’s search filter on patient preferences,³¹ Harbor’s filter on patient issues,³² and sensitive search strategies on psychosocial and behavioral issues as used in systematic reviews to underpin guidelines of the UK National Institute for Health and Clinical Excellence (NICE).³³ Nevertheless, differences in domain coverage remained. Moreover, none of the above strategies had been developed by means of word frequency analysis or accompanied by evaluation and validation data. Finally, all were search strategies prioritizing sensitivity, and as

a result were rather lengthy (eg, Harbor’s filter on patient issues consisted of 142 lines).³² The word frequency analysis study reported here thus adds to the resources for bibliographic database searching in this increasingly important aspect of healthcare.

Objective approaches to search filter development using word frequency analysis

Objective approaches to search filter development aspire to be transparent, replicable, and include a minimum of subjective decisions. Objective approaches using word frequency analysis were pioneered more than a decade ago by Boynton *et al*³⁴ on the titles, abstracts, and subject headings of 288 systematic reviews. White *et al*³⁵ enhanced this study design by analyzing MeSH terms and employing discriminant analysis to select and combine words into search strategies. Subsequent research has drawn extensively on these two studies, including in terms of word frequency analysis software used—the ListServe function of the Blackwell bibliographic software Idealist and the WordStat module of the SimStat statistical analysis software.

The majority of work within the field aimed to develop ‘methodological filters’—such as for systematic reviews,^{34 35} randomized controlled trials,^{36 37} and diagnostic studies.^{38 39} Research with a substantive thematic focus has been limited and less successful. For example, Wentz *et al* were unable to devise

Box 2 What are health-related values?

A. Understanding of values as per initial brief (for scoping searches performed during background work)

The aim of the scoping searches was: (a) to identify research on patients' and other stakeholders' values (understood broadly to include perceptions, preferences, experiences, wishes, concerns, points of view, voices, stories, etc); (b) on the way these values affect health-related choices, decisions, and behaviors; (c) on the differences between patients' and other stakeholders' values, in particular between patients and health professionals; (d) on the diversity and disagreements about values; and (e) on the often implicit values that affect professional models of care—for instance, values in diagnosis.

At a more abstract level, we were looking for studies that explored or illustrated different visions of what is good, right, needed, important, etc and different ways in which a positive or negative sign was assigned with regard to health, illness and healthcare. This also included research addressing complex issues with some values elements (eg, organizational change), regardless of whether the presence of values was openly recognized or not.

B. Grouping of values after initial searches—based on the intersection of expectations based on initial brief and issues found in the literature. Used to anchor judgments of saturation:

- ▶ The value-laden way of apprehending (perceiving, understanding, thinking about things, etc) ('perception', 'attitude', 'perspective')
- ▶ Values as demonstrated through choices, preferences, decisions, etc ('choice', 'goal', 'wish')
- ▶ Values as individually specific ('identity', 'personal', 'subjective')
- ▶ Values of the immediate social system of the individual ('carer', 'friend', 'social support')
- ▶ Values of the wider social system ('culture', 'social', 'public')
- ▶ The existential, spiritual and religious dimension of values ('existential', 'meaning', 'spiritual')
- ▶ The moral, ethical and legal dimensions of values ('duty', 'obligation', 'dilemma')
- ▶ Values of patients in their 'patient persona' ('advocacy', 'patient groups', 'empower')
- ▶ Values of the various health professions as elicited through their interaction ('interdisciplinary', 'team working')
- ▶ Values in science ('paradigm', 'model')
- ▶ Manifestations and transformations of values in the experience of illness and recovery ('journey', 'story', 'experiences')
- ▶ Communicating and failing to communicate values ('conflict', 'listen', 'voice')
- ▶ Values associated with the interaction with the healthcare system ('barrier', 'satisfaction', 'good doctor')

C. Understanding of values as per inter-coder reliability exercise. Instructions were kept close to (A) rather than (B) above. The aim was to ascertain to what extent intuitive judgments, without the specification of detailed inclusion—exclusion criteria, would result in acceptable degree of inter-coder agreement:

'A true positive is any reference for a publication that can be expected to give us some knowledge and understanding of health-related values—of different visions of what is good, right, needed, important ..., of different ways in which a positive or negative sign can be assigned with regard to health, illness, and healthcare. This includes issues such as patients', carers', health professionals', researchers' and other stakeholders' values, attitudes, perceptions, preferences, perspective, beliefs, priorities, wishes, etc, but need not be restricted to them (and is definitely not ethics only). There are a number of gray areas, and the idea is to see whether your assignments, but also your hesitations, are similar or different to mine. Please be inclusive in your selection rather than narrowly focused. The idea at this stage is to be as non-prescriptive and non-restrictive as possible.'

D. The following is our most recent taxonomy that aims to improve on the theoretical coherence and heuristic potential of (B). This is still a 'working hypothesis'. It suggests that publications on health-related values are ones that explore or illustrate:

- ▶ *processes of evaluation*- any process of assigning a positive or negative sign to a phenomenon (eg, *processes of liking and disliking, of choosing and making decisions*) in the context of health, illness and healthcare;
- ▶ any substantivisation, realization, or precondition for such processes in the form of *values, value-laden phenomena, value-laden configurations, situations, agents, etc* including:
 - specific outcomes of and prerequisites for choosing and making decisions—such as *preferences, priorities, choices*;
 - highly abstract ends in terms of which evaluations are made—such as *good, right, wrong, value, quality*;
 - specific instantiations of those highly abstract ends—such as *care, autonomy, support*; or *maleficence, negligence, failure*;
 - experiences, sensations, perceptions, etc that are evaluatively defined, have multiple value-laden consequences and may strongly color other evaluations—such as experiences and sensations of *pain, despair, loneliness*; or of *recovery, well-being, self-efficacy*;
 - occurrences and behaviors that are evaluatively defined and have multiple value-laden consequences—such as *falling ill, bereavement, self-harming*; or *recovery, self-management, coping behaviors*;
 - experiences and situations that are defined by either differences and discrepancy, or agreement and congruence about evaluations, values, value-laden phenomena, etc—such as *disagreement, conflict, consensus*;
 - broad contexts, situations, arenas in which those evaluations, values, or agreement and differences about them are formed, experienced, realized, contested, and modified, which can be more persistent and pervasive (such as *culture, religious beliefs, lifestyle, socioeconomic status*) or more specific and context-limited (such as *instances of communication, patient-health professional encounters, illness experiences, public health interventions*);
 - actors who embody certain values or are a common target of evaluations (such as *vulnerable populations, a good doctor, compliant and non-compliant patients*).

search strategies combining acceptable sensitivity and precision for controlled studies on road safety interventions.⁴⁰ Goss *et al*, working in the context of alcohol-impaired driving interventions, suggested that, with multidisciplinary topics and heterogeneous vocabulary, word frequency analysis findings may be a useful starting point, but substantial subjective input may be needed.⁴¹ Sladek *et al*, in a study aiming to improve a palliative care filter, concluded that future improvements depended on better identifying how clinicians and researchers conceptualized such care.⁴² Therefore, the present study was carried out in the context of limited work on the objective development of 'thematic' search filters.

OBJECTIVES

This study aimed to:

1. Generate a broad pool of keywords representing values concepts, along with data on their sensitivity and precision. This was intended to serve primarily the needs of researchers and library and information specialists involved in search strategy design.
2. Develop a brief search filter (≤ 20 lines) of high precision ($\geq 67\%$) and acceptable sensitivity ($\geq 67\%$) which can be used to perform broad-scoping searches for publications on values in generic medical and health databases (eg, Medline, Embase, or Cinahl) and across a range of topics (eg, health conditions, health settings, health interactions, etc). This was intended primarily for health professionals, educators, health service planners, policy makers, and other users interested in an overall picture of the values associated with a particular health condition or context.

METHODS AND SOURCES

Derivation datasets

Diabetes, obesity, dementia, and schizophrenia were selected as conditions: (1) having a significant physical and psychosocial impact on patients and carers; (2) placing a substantial economic and organizational burden on health and social services; and (3) addressed by new policy documents, consultations, and/or intense media debates in the UK during the study design stage (2004–2005). We used as derivation datasets the most robustly compiled and richest collections of citations on these topics from a set of collections on values materials we had accumulated over a period of 4 years. Table 1 outlines the main elements of this background work (see online supplementary file 1, section 1). Online supplementary file 1, section 2 describes in detail the process of compiling the datasets. Briefly, the datasets contained both 'true' and 'false positive' citations ('true positives' were citations classified as likely to represent values publications). They covered the period January 2004–December 2006 and were obtained through Medline searches using a set of values search terms shown by previous work to be sensitive and/or precise. The process of compiling the derivation datasets was terminated once saturation was reached and the number of duplicates began to exceed one-quarter of the entries. Tests were performed to assess the likelihood of bias within the sample. This was ascertained as low.

A randomly selected 20% of the citations from all derivation datasets were subjected to two-phase inter-coder testing with regard to their true/false positive status (performed by MP and PS, instructions in box 1). The first phase showed an average of 24.7% disagreement. In the second phase, the citations with conflicting assignments from the first phase were recoded independently by each reviewer. The average of the remaining

disagreements was 11.6%. A final assignment was reached through consensus. It was judged that, in view of the complexity of the concept of health-related values, this level of uncertainty in the classification into true and false positives was admissible.

Word frequency analysis for text terms

For each topic, two *txt* files—one of true positive and one of false positive citations—were prepared containing titles, abstracts, and MeSH terms. All *txt* files were generated by first creating a customized Endnote output style. Normally, these output styles are used to generate customized bibliographies (function available through Edit/Output Styles/New Style). In our case the output style and, through it, each reference in the 'bibliography', contained the title of the paper, its abstract and MeSH terms. Those 'bibliographies' were generated in a Word document (through *Insert selected citations*) and then saved as *txt* files.

For each topic, the aim of the analysis was to identify the most frequent and precise values words appearing in the true positive citations. Frequencies were obtained by making a 'full concordance' of a true positives file using the Concordance® software⁴³ (see box 1). The calculation of precision statistics required an extended sequence of steps so that it could be automated. The difficulty lay in enabling the matching of words that were shared between the true and false positives. The process is described in online supplementary file 1, section 3. It was performed on words with a frequency ≥ 10 in order to make the datasets more manageable and reduce the likelihood of errors.

Word frequency analysis of MeSH terms

Txt files containing MeSH terms from the true positives and false positive citations were prepared for each of the topics (eight files in total). These included an identification number for each citation and the MeSH terms used to index it. The customized files were generated as per the approach described above.

The *txt* files were then imported into SPSS. The Concordance® software could not be used, as it only counts single or graphically connected words (eg, with a hyphen or underscore) while a substantial proportion of the MeSH terms are phrases or compound words. Frequency counts were produced for the MeSH terms in both the true positive and false positive datasets. For the most frequent MeSH in the true positive datasets, the respective frequencies in the false positive datasets were identified (manually), and precision statistics were computed.

Design of a brief values filter

To enable the design of a brief values filter, the word frequency analysis had to identify a sufficient number of words of high precision and moderate to high sensitivity that were shared by at least three topics. A more permissive inclusion criterion was chosen (ie, words shared by three rather than four conditions) so as to capture values concerns that are important in a wide range of conditions but may have a low profile in a particular one. For instance, background work suggested that stigma may be a secondary concern in the case of diabetes, but is highly evident in obesity, schizophrenia, and, to an extent, dementia.

Validation work

Internal validation of the brief values filter was performed on a combined dataset for all derivation topics. The unique PubMed identifiers (PMID) of the true positives were combined into

a search string. The same procedure was employed for the false positives. This permitted us to reproduce the derivation datasets as PubMed searches. A separate search was performed using the brief values filter. This was first combined with an AND with the search representing the true positives dataset. The overlap in records reflected the sensitivity of the brief filter in the combined derivation dataset. The search using the brief filter was then combined with an AND with the search representing the false positives dataset. The overlap in records represented the number of false positives picked by the filter. The specificity and precision of the brief values filter in the combined derivation dataset were calculated as follows: specificity = (all false positives – false positives picked by filter)/all false positives; precision = true positives picked by filter/(true positives picked by filter + false positives picked by filter).

External validation of the brief values filter was performed on citations on dentistry and food allergies, which were selected as two topics contrasting with the derivation ones. Sample size estimates were based on generating 95% CIs of width 10% for sensitivity and precision of the search filter using the method of Buderer.⁴⁴ The validation datasets were obtained by performing searches in PubMed for Dentistry [MeSH] and Food Hypersensitivity [MeSH]. Citations were reviewed up to respective point estimates for sample size and classified into cases and non-cases. The unique PMID identifiers for different groups of citations were combined into search strings. Sensitivity, specificity, and precision statistics were calculated as described above.

RESULTS

Size and composition of the derivation datasets

Table 1 describes numerically the datasets subjected to word frequency analysis. A total of 4440 citations were analyzed, consisting of over 1 110 000 words.

Precision statistics were calculated for 1821 text words in the case of diabetes, 1591 in the case of obesity, 1701 for dementia, and 1541 for schizophrenia (all words from the respective true positives dataset which had a frequency ≥ 10). For each of the derivation topics, the 50 most frequent and precise text words and MeSH (also referred to as the '50 best' sets of text words or MeSH) are reported below.

Overlap between topics

Across the four '50 best' sets of text words, there was a total of 144 different words. Thirty-seven (25.7%) were shared between two or more conditions. Of these, four (2.8%) were shared by all four conditions, and 15 (10.4%) by at least three conditions. Dementia and schizophrenia shared the largest number of frequent and precise text words for values (17), and obesity and schizophrenia the smallest (7).

Across the four '50 best' sets of MeSH terms, there was a total of 124 different terms. Forty-five MeSH terms (36.3%) were shared between two or more conditions. Of these, nine (7.3%) were shared by all conditions, and 22 (17.7%) by at least three

conditions. Diabetes and schizophrenia were shown to have the largest number of frequent and precise MeSH terms for values in common (24), and obesity and dementia the smallest (13).

Best performing text words for values

Table 3 (online supplementary file 2) lists the 50 best performing text words for values for each of the derivation conditions, along with their frequency and precision statistics (see table 2 for an initial segment). These have been drawn from a list of text words with a frequency ≥ 20 and precision ≥ 0.90 . Frequencies tend to reflect the number of uses of a word rather than number of abstracts: if the same word is used twice in an abstract, Concordance® counts two uses. Data cleaning was performed for text words with a 100% precision, and in those cases the reported frequencies represent number of abstracts. Same-root words are listed separately, as they may have very different precision (eg, in diabetes, 'perceptions' had a 100% precision, while 'perception' generated false positives in more than one-third of cases). Counts also include MeSH terms or parts of MeSH terms, as, in this element of the analysis, MeSH terms were treated as any other text word. MeSH terms were included in the analysis of text words, as, in PubMed, a text word search will search indexing fields by default.

Best performing MeSH terms

Table 4 (online supplementary file 2) lists, for each condition, the 50 best performing MeSH terms along with their frequency and precision statistics (see table 3 for an initial segment). MeSH terms were excluded if they referred to the country where the study was carried out. The reported terms have been drawn from a list of MeSH terms with precision ≥ 0.90 .

Brief values filter

Nineteen text words and seven MeSH terms were identified that had simultaneously a very high precision ($\geq 90\%$) and high sensitivity (≥ 30 occurrences for text words and ≥ 20 occurrences for MeSH terms) across at least three of the derivation topics. Of the 26 terms identified, five contained the word 'attitude' ('attitude'; 'attitudes'; Attitude of Health Personnel [MeSH]; Attitude to Health [MeSH]; Health Knowledge, Attitudes and Practice [MeSH]). These were brought together under 'attitude*'. As a result, 21 unique terms were used to compile the brief values filter (see box 3).

Performance of the brief values filter in internal and external validation

The brief values filter was able to capture 76.8% of the citations in the combined derivation dataset (internal validation on diabetes, obesity, dementia, and schizophrenia). It had a specificity of 85.5% and precision of 86.8%.

In the external validation, the dataset on Food Hypersensitivity consisted of 1214 citations (June 2008–July 2010), and

Table 1 Size and composition of datasets subjected to word frequency analysis

	True positives			False positives		
	No of citations	Words (including repetitions)	Unique words	No of citations	Words (including repetitions)	Unique words
Diabetes	668	163 087	12 142	536	146 811	14 211
Obesity	582	139 970	11 264	488	129 038	12 743
Dementia	639	149 140	11 432	517	132 040	12 796
Schizophrenia	560	134 955	10 576	450	115 250	10 490
	2449	587 152		1991	523 139	

Total number of citations, 4440. Total number of words, 1 110 291.

Table 2 First 15 best performing text words for values (see online supplementary file 2 for full table)

Diabetes			Obesity			Dementia			Schizophrenia			
Word	Frequency	Precision	Word	Frequency	Precision	Word	Frequency	Precision	Word	Frequency	Precision	
1	Psychology	312	0.97	Psychology	306	0.92	Caregivers	583	0.95	Life	376	0.91
2	Social	253	0.95	Image	231	0.95	Family	437	0.93	Attitude	178	0.97
3	Satisfaction	201	0.97	Attitudes	220	0.96	Life	353	0.90	Attitudes	154	0.98
4	Self	188	0.95	Self	134	0.92	Support	324	0.90	Coping	135	0.97
5	Questionnaires	188	0.92	Satisfaction	125	0.95	Quality	309	0.93	Qol	123	1.00
6	Psychological	184	0.97	Attitude	123	0.95	Caregiver	283	0.95	Intervention	118	0.91
7	Interviews	175	0.98	Perceived	121	0.91	Services	210	0.93	Relations	110	0.90
8	Attitude	173	0.99	Dissatisfaction	96	0.99	Burden	178	0.94	Public	97	0.91
9	Self-management	164	0.94	Counseling	96	0.91	Attitude	175	0.97	Interview	92	0.91
10	Patients'	148	0.91	Barriers	89	0.93	Staff	170	0.96	Personal	90	0.90
11	Attitudes	141	0.96	Self-esteem	87	0.99	Relations	168	0.91	Nursing	89	0.97
12	Qualitative	134	0.99	Binge	86	0.95	Caregiving	129	0.96	Adherence	87	0.98
13	Relations	132	0.99	Perceptions	82	1.00	Interviews	128	0.93	Psychosocial	86	0.92
14	Nursing	122	0.97	Beliefs	77	0.99	Program	120	0.92	Burden	83	0.94
15	Educational	120	0.90	Interviews	66	0.93	Caregivers/*	112	0.97	Toward	81	0.90

*Here caregivers/ is part of an indexing string, eg, caregivers/ psychology. The software has treated the word and the slash as a single word and separate from caregivers heading the list.

the one on Dentistry consisted of 1062 citations (Mar 2010–July 2010). Table 4 summarizes the validation data.

A number of 'subjective' additions to the objectively derived filter were tested (selected to improve consistency, eg, spelling variants, or on the basis of exploration of excluded cases). They resulted in variable improvement in sensitivity and associated decrease in precision and specificity. For instance, adding 'quality of life' (in addition to QOL and Quality of Life [MeSH]) raised sensitivity by 1.1 percentage point, to 77.9% (1862/2390) in the combined derivation dataset. Specificity and precision fell, respectively, by 1.3 and 0.9 percentage points. In the case of Dentistry, adding 'esthetic*' raised sensitivity by over five percentage points, to 52.9% (128/242). The drops in specificity

and precision were, however, similarly substantial, respectively 4.5 and 14.9 percentage points.

DISCUSSION

Summary of main findings

This paper described the design and findings of a word frequency analysis study of over 4400 citations (titles, abstracts, and indexing terms) for publications on 'health-related values'. The latter were construed as anything positively or negatively weighted that could affect health-related decisions and actions, to include phenomena such as attitudes, perceptions, perspectives, priorities, beliefs, preferences, choices, wishes, rights, experiences, satisfaction, quality of life, etc. Research on such

Table 3 First 15 best performing MeSH terms for values (see online supplementary file 2 for full table)

Diabetes			Obesity			Dementia			Schizophrenia			
MeSH	Frequency	Precision	MeSH	Frequency	Precision	MeSH	Frequency	Precision	MeSH	Frequency	Precision	
1	Questionnaires	148	0.91	Body image	82	1.00	Caregivers	187	0.94	Quality of life	87	0.95
2	Patient education as topic	121	0.90	Self concept	76	1.00	Quality of life	83	0.97	Attitude to health	84	0.98
3	Self care	116	0.94	Attitude to health	62	0.94	Attitude of health personnel	74	0.96	Adaptation, psychological	68	0.94
4	Attitude to health	114	0.98	Quality of life	48	0.94	Adaptation, psychological	73	0.99	Social support	47	0.98
5	Health knowledge, attitudes, practice	86	0.93	Prejudice	35	1.00	Social support	68	0.99	Attitude of health personnel	36	0.97
6	Quality of life	79	0.95	Attitude of health personnel	32	0.97	Attitude to health	68	0.97	Family	34	0.94
7	Patient satisfaction	63	0.95	Motivation	29	1.00	Decision making	54	0.93	Stereotyping	33	1.00
8	Adaptation, psychological	62	1.00	Stereotyping	23	1.00	Family	53	0.96	Caregivers	32	1.00
9	Physician–patient relations	47	0.98	Nurse's role	22	0.96	Cost of illness	49	0.94	Cost of illness	32	0.94
10	Social support	43	1.00	Self efficacy	22	0.92	Home nursing	34	0.94	Psychiatry	29	0.91
11	Attitude of health personnel	37	1.00	Focus groups	19	1.00	Interviews as topic	33	0.92	Culture	25	0.96
12	Interviews as topic	36	0.97	Social support	18	1.00	Nursing methodology research	32	0.97	Cross-cultural comparison	23	0.96
13	Focus groups	30	1.00	Adaptation, psychological	15	0.94	Health knowledge, attitudes, practice	31	0.97	Health knowledge, attitudes, practice	22	1.00
14	Patient-centered care	29	0.91	Personal satisfaction	15	0.94	Patient-centered care	31	0.97	Psychiatric nursing	22	1.00
15	Qualitative research	28	1.00	Interpersonal relations	14	1.00	Mental competency	31	0.94	Patient education as topic	21	1.00

Box 3 Brief high precision search filter for scoping searches for publications on health-related values in Medline

1. Attitude* (tw)
2. Perceptions (tw)
3. Qualitative (tw)
4. Coping (tw)
5. Counseling (tw)
6. Cultural (tw)
7. Ethics (tw)
8. Experiences (tw)
9. Interviews (tw)
10. Perceived (tw)
11. Personal (tw)
12. Professionals (tw)
13. QOL (tw) OR Quality of Life (mh)
14. Relations (tw)
15. Respondents (tw)
16. Satisfaction (tw)
17. Staff (tw)
18. Well-being (tw)
19. Adaptation, Psychological (mh)
20. Nurse's Role (mh)
21. Social Support (mh)
22. OR/1-21

This search filter was developed on the basis of Medline citations. Its adaptation to other databases requires further research.

The words in lines 1–18 are text words, with the exception of 'Quality of Life', which was added to line 13 for thematic consistency (it was itself a high frequency, high precision MeSH). The remaining words are MeSH terms. Each of the included search terms had a precision $\geq 90\%$ and a frequency ≥ 30 (for text words) or ≥ 20 (for MeSH terms) in at least three of the four derivation conditions. A higher cut-off point was chosen for the frequency of text words as it reflected occurrences of the word rather than the number of abstracts in which they appeared.

We suggest that text words are searched without being mapped on to thesaurus terms (for the PubMed interface, this is achieved by adding the tag (tw), as above; in Ovid, this means not selecting the mapping option). This is because mapping may diminish the precision of text words as calculated on the basis of the word frequency analysis in this study. The final line indicates that the terms are to be combined with the Boolean operator OR. The resulting output can then be combined (with an AND) with topic-specific terms (eg, of a health condition; healthcare setting—eg, primary care, hospital, public health; type of stakeholder—patients, professionals, carers, specific ethnic groups, etc) to obtain an overview of the values research concerning that topic.

issues is scattered and may be difficult to locate, yet is crucial to understanding and improving clinical practice, health policy, and healthcare delivery.

Only 2.8% of the best performing text words and 7.3% of the respective MeSH terms were shared by all four topics studied. It is unclear whether this indicates different vocabulary and research traditions or actual differences in key values concerns. As far as pair-wise comparisons of shared values terms are concerned, it was surprising that diabetes and schizophrenia had the largest number of frequent and precise MeSH terms in common. Following the screening process, we were of the impression that in comparison with diabetes and schizophrenia, research on obesity and dementia was more 'paternalistic' and less interested in the voices of patients. If this impression is adequate, then research on obesity and dementia may have

lower frequency of certain values concepts for contingent reasons, including through reflecting broader societal bias and prejudice. Alternatively, the values profiles of diabetes and schizophrenia may be 'essentially' more similar to one another. Values profiles may give grounds for drawing new lines of similarity and difference between health conditions, with implications for the clinical and public health management of the newly formed groups.

In spite of the small number of terms shared by all four topics, it was possible to design a brief Medline search filter for scoping searches for values publications that appears generalizable across topics. Its performance parameters were high: in the derivation dataset, sensitivity was 76.8% (aimed for 67%) and precision 86.8% (aimed for 67%). When validated in new datasets, the performance parameters were still satisfactory (food

Table 4 Performance statistics for the brief values filter

	Derivation dataset (combined diabetes, obesity, dementia, and schizophrenia)	Food hypersensitivity (1214 citations, true positives 8.0%)	Dentistry (1062 citations, true positives 22.8%)
Sensitivity (true positives picked/all true positives)	76.8% (1836/2390)	70.1% (68/97)	47.1% (114/242)
Specificity (false positives avoided/all false positives)	85.5% (1654/1934)	96.5% (1078/1117)	97.1% (796/820)
Precision (true positives picked/all picked)	86.8% (1836/2116)	63.6% (68/107)	82.6% (114/138)

hypersensitivity—sensitivity of 70.1%, precision of 63.6%; dentistry—sensitivity of 47.1%, precision of 82.6%). It was also observed that minor additions to the original filter may noticeably affect its performance in new areas—for example, in the case of Dentistry, adding ‘esthetic*’ increased its sensitivity to 52.8% but decreased its precision to 67.7%. This suggests that the brief values filter can be applied productively to other topics in Medline, but that there is instability in its performance. To what extent this follows from the nature of values research and the difficulties of it being captured, the differences of values across conditions or from an insufficient sample size is a question requiring further research.

There may be something particular about the nature of values research that accounts for the generally high precision and sensitivity of the brief values filter. Values studies tend to discuss a cluster of values issues rather than only one such issue, and the associated abstracts tend to contain a number of values terms. Therefore, values citations can be accessed through a number of entry points, and a single values term will most likely lead to a network of topics and concepts beyond the ones narrowly associated with the concept it encodes. Apparently, there are concepts that belong to a large number of networks of values issues. At least some of the words in the brief values filter are likely to represent such concepts.

Strengths

In comparison with earlier research using word frequency analysis for search filter design, this study used a uniquely large derivation dataset. It consisted of over 4400 citations and over one million words. In the previous research identified, the range was between 61 and 1729 records^{34 35}; the mean was 396 records. However, in terminating the process of compiling the derivation datasets, we relied on a perception of saturation of main topics, concepts, issues, and messages. This was complemented by a rule of thumb from linguistics that a corpus of at least half a million words is needed to produce reliable findings. Further work is required to propose more objective standards for the size of derivation datasets when complex and multifaceted topics, such as health-related values, are explored.

This study was also unusual in having analyzed MeSH terms that were compound words or phrases. Few studies appear to have done this so far,^{35 42 45} as word frequency analysis software cannot count graphically disconnected semantic units. Here, a combination of the functionalities of Endnote and SPSS was used to perform the required analysis.

Most importantly, this study provides sound, evidence-based tools for identifying publications on health-related values as well as a starting point from which those tools can be further developed. The broad pool of search terms and the brief search filter will enable the efficient compilation of sizeable datasets for further analysis, which was a major obstacle in this study.

Limitations

The main limitation of this study is that it covers a limited period, limited number of topics, a single database, and a single approach to compiling its corpus of citations. Furthermore, it lacked well-articulated criteria for distinguishing cases (of values publications) from non-cases. The criteria used may have been overly inclusive. If stricter criteria are applied in future research, the precision of terms from the broad pool as well as of the brief values filter may diminish. Developing unambiguous inclusion–exclusion criteria requires further conceptualization of health-related values. This study did not have the capacity for both theoretical and empirical work of adequate depth.

Currently, we are examining boundary cases (work in progress, part of a manual on searching for values publications), and this is showing promise toward improved specification of inclusion–exclusion criteria for values publications. Nevertheless, ‘values’ is a topic where tight inclusion–exclusion criteria may always be problematic, where the validity and reliability of assignments may be in constant tension. This is a substantial challenge in the field of objective search strategy design, where transparency, replication, and minimization of subjective decisions are paramount.

Another important limitation is that the reported frequencies for text words did not necessarily reflect the number of abstracts in which a word appears. It is the latter that would be a true indicator of sensitivity. Data cleaning was performed on text words that had a 100% precision, but not for the remainder. The manual procedure was rather time-consuming and there was no identifiable function in the word frequency analysis software that allowed automation. No such margin of error is present in the case of MeSH terms, as these are not repeated in one and the same citation.

Finally, a more fundamental concern can be raised that applies not only to this study but to the very field of designing objective search strategies on the basis of word frequency analysis. The field may be lagging behind advances in medical informatics. Some of the methods applied in this study may need to be superseded, including methods that have improved on standard practices of objective search strategy design. For instance, post-study feedback suggested that difficulties of handling compound words may have been more efficiently overcome by Unix/Linux scripts and simple programming in a standard language (such as C or Perl). It remains to be ascertained if the approaches offered here provide a lengthier but equally effective equivalent to programming procedures, with the advantage of being accessible to researchers with no programming skills, or whether they have substantial disadvantages. If the latter proves to be the case, then having a health informatician will not only be desirable and advantageous but required for teams working on search strategy development.

Directions for future research

As far as search strategies for values publications are concerned, further work is needed that uses different approaches to compiling derivation datasets and targets different databases, time periods, and topics. Theoretical and conceptual research is also required to inform decisions about scope of searches and inclusion/exclusion criteria. More broadly, relative to the context of objective methods for search strategy design, priority questions include what the comparative strengths and weaknesses of various approaches to compiling derivation datasets are, what makes adequate sizes of derivation and validation datasets, and how to overcome technical challenges to performing helpful analyses. The latter may be fed back to developers of database interfaces and word frequency analysis software. For instance, in order to analyze compound MeSH terms, we instituted a multi-step procedure of generating, through Endnote output styles, files that contained only an indexing number and MeSH terms, reading the files in SPSS, automating the separation of subheadings, performing frequency counts, etc. Much of this complexity would have been unnecessary had the graphical convention for compound MeSH terms been different—for example, if the elements of the compound term were connected through hyphens. A wider range of formats in which data can be downloaded from databases would also contribute to simpler study designs.

Ultimately, however, the aim of this research is to facilitate utilization and uptake of research on health-related values—in actual applications, as in clinical practice, policy development and medical education, but also in research cross-fertilization. It is the needs of users within these fields that should, first and foremost, shape directions for future research on developing search strategies for identifying materials on health-related values.

CONCLUSIONS

Word frequency analysis has shown promising results and huge potential in the development of search strategies for identifying publications on health-related values. Other ‘diffuse’ topics,⁴⁶ such as change (both in healthcare organizations and of health behaviors), communication, social support, learning, and teaching may also lend themselves to effective exploration for the purposes of search strategy design through these or similar techniques from the field of the health information sciences.

Acknowledgments We thank the Laces Trust and the West Midlands Deanery, and personally Mr Peter Beesley and Professor Steve Field, for funding a number of stages of this project. We also thank Dr Andrew Shanks and Jacqui Cox for their essential contribution to the early days of the project—respectively for his applied linguistics advice and initiating the feasibility word frequency analysis, and for her library sciences expertise and support in testing and refining the initial list of search terms. We would also like to acknowledge Bastiaan Brak and Abigail Lee for their invaluable help with data entry, members of the steering group for their advice and facilitation of contacts, Dr Harbinder Sandhu for contributing to the feasibility inter-coder reliability study, and Dr Charlotte Price for her statistical advice. We are also indebted to Diabetes UK for carrying out a search for values publications on their in-house resources. Finally, but very importantly, we would like to thank our five reviewers for their detailed and often challenging feedback, which has completely transformed this paper.

Funding The Laces Trust; West Midlands Deanery; Warwick Medical School.

Competing interests None.

Contributors MP contributed to the general framework for the study design, developed and tested out the specific steps within the study design, carried out the scoping searches, compiled the analysis datasets, carried out the data analysis, and drafted and edited the manuscript. PS contributed to the refinement of specific steps within the study design, acted as coder in the inter-coder reliability exercise reported in the paper, and commented on and edited numerous versions of the paper. KWMM provided theoretical input into conceptualizing and operationalizing the concept of health-related values, contributed to the general framework for the study design, provided feedback on specific steps within it, acted as coder in a feasibility inter-coder reliability study, contributed to the drafting of the background sections, and commented on and edited numerous versions of the paper. JD contributed to the general framework for the study design, provided feedback on specific steps within it, acted as coder in a feasibility inter-coder reliability study, and commented on and edited numerous versions of the paper.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Department of Health.** *The Expert Patient: A New Approach to Chronic Disease Management for the 21st Century.* London: Crown, 2001. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4018578.pdf (accessed 12 Mar 2011).
2. **Department of Health with Royal College of Psychiatrists and National Institute for Mental Health in England.** *New Ways of Working for Psychiatrists: Enhancing Effective, Person-Centred Services Through New Ways of Working In Multidisciplinary and Multi-Agency Contexts.* London: Crown, 2005. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4122343.pdf (accessed 12 Mar 2011).
3. **Royal College of Psychiatrists.** Partners in care campaign. <http://www.rcpsych.ac.uk/campaigns/pinc/> (accessed 12 Mar 2011).
4. **Religion or belief.** *A Practical Guide for the NHS.* London: Department of Health, 2009. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_093132.pdf.
5. **Department of Health (Fish, Julie for the DH Sexual Orientation and Gender Identity Advisory Group).** *Reducing Health Inequalities for Lesbian, Gay, Bisexual and Trans People - Briefings for Health and Social Care Staff.* London: Crown, 2007. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_078347 (accessed 12 Mar 2011).
6. **Department of Health.** *Putting People First: A Shared Vision and Commitment to the Transformation of Adult Social Care.* London: Crown, 2008. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_081118 (accessed 12 Mar 2011).
7. **Department of Health.** *Valuing People: A New Strategy for Learning Disability for the 21st Century.* London: Crown, 2001.
8. **Cleland JGF, Ekman I.** Enlisting the help of the largest health care workforce - patients. *JAMA* 2010;**304**:1383–4.
9. **World Health Organization.** *Patient Self-Management Booklet,* 2006. http://www.who.int/hiv/pub/imai/patient_self/en/index.html (accessed 12 Mar 2011).
10. **Department of Health.** *Creating a Patient-Led NHS: Delivering the NHS Improvement Plan.* London: Crown, 2005. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4106506 (accessed 12 Mar 2011).
11. **Department of Health.** *Commissioning a Patient-Led NHS.* London: Crown, 2005. <http://www.dh.gov.uk/assetRoot/04/11/67/17/04116717.pdf>.
12. **Department of Health.** *High Quality Care For All: NHS Next Stage Review Final Report.* London: Crown, 2008. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_085825.
13. **Department of Health.** *Our Health, Our Care, Our Say: A New Direction for Community Services (White Paper).* London: Crown, 2006. <http://webarchive.nationalarchives.gov.uk/+www.dh.gov.uk/en/Healthcare/OurHealthourCareourSay/index.htm>.
14. **Black N, Jenkinson C.** Measuring patients' experiences and outcomes. *BMJ* 2009;**339**:2495.
15. **Measuring Effectiveness and Cost-Effectiveness: The QALY.** <http://www.nice.org.uk/newsroom/features/measuringeffectivenessandcosteffectiveness/theqaly.jsp>.
16. **Department of Health.** *Putting People at the Heart of Care: The Vision for Public and Patient Experience and Engagement in Health and Social Care.* London: Crown, 2009. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_106038 (accessed 12 Mar 2011).
17. **World Health Organization.** *Patients for Patient Safety - Statement of Case.* http://www.who.int/patientsafety/patients_for_patient/statement/en/index.html (accessed 12 Mar 2011).
18. **Ahmed SM, Palermo A-GS.** Community engagement in research: frameworks for education and peer review. *Am J Public Health* 2010;**100**:1380–7.
19. **Colombo A, Bendelow G, Fulford KWM, et al.** Evaluating the influence of implicit models of mental disorder on processes of shared decision making within community-based multidisciplinary teams. *Soc Sci Med* 2003;**56**:1557–70.
20. **Wong S, Wilczynski N, Haynes R; the Hedges Team.** Developing optimal search strategies for detecting clinically relevant qualitative studies in MEDLINE. *Stud Health Technol Inform* 2004;**107**(Pt 1):311–16.
21. **Jenkins M.** Evaluation of methodological search filters—a review. *Heath Info Libr J* 2004;**21**:148–63.
22. **Oxford English Dictionary Online.** <http://dictionary.oed.com/>.
23. **Von Wright GH.** *The Varieties of Goodness.* London: Routledge & Kegan Paul, 1963.
24. **Hare RM.** *The Language of Morals.* Oxford: Oxford University Press, 1952.
25. **Hunink M, Glasziou P, Siegel J, et al.** *Decision Making in Health and Medicine: Integrating Evidence and Values.* Cambridge: Cambridge University Press, 2001.
26. **Brown M, Brown G, Sharma S.** *Evidence-Based to Value-Based Medicine.* Chicago: American Medical Association Press, 2005.
27. **Kahn T, Ninomiya H.** Changing vocabularies: a guide to help bioethics searchers find relevant literature in national library of medicine databases using medical subject headings [MeSH] indexing vocabulary. *Kennedy Inst Ethics J* 2003;**13**:257–311.
28. **Shaw RL, Booth A, Sutton AJ, et al.** Finding qualitative research: an evaluation of search strategies. *BMC Med Res Methodol* 2004;**4**:5.
29. **Flemming K, Briggs M.** Electronic searching to locate qualitative research: evaluation of three strategies. *J Adv Nurs* 2007;**57**:95–100.
30. **Cochrane Consumers and Communication Group on the Cochrane Library website.** http://www.mrw.interscience.wiley.com/cochrane/cochrane_clsystrev_crglist_fs.html (accessed 12 Mar 2011).
31. **National Primary Care Research and Development Centre** “Patient Preferences in Primary Care” Filter. <http://www.york.ac.uk/inst/crd/intertasc/public1.htm> (accessed 12 Mar 2011).
32. **“Patient Issues” Filter.** <http://www.york.ac.uk/inst/crd/intertasc/public2.htm> (accessed 12 Mar 2011).
33. **National Collaborating Centre for Mental Health.** Appendix 13a (Review protocols: CD-ROM) to National Collaborating Centre for Mental Health: A NICE-SCIE Guideline on supporting people with dementia and their carers in health and social care, 2007. <http://www.scie.org.uk/publications/misc/dementia/dementia-fullguideline.pdf>.
34. **Boynton J, Glanville J, McDaid D, et al.** Identifying systematic reviews in MEDLINE: developing an objective approach to search strategy design. *J Inform Sci* 1998;**24**:137–57.
35. **White VJ, Glanville JM, Lefebvre C, et al.** A statistical approach to designing search filters to find systematic reviews: objectivity enhances accuracy. *J Inform Sci* 2001;**27**:357–70.
36. **Glanville J, Lefebvre C, Miles J, et al.** How to identify randomized controlled trials in MEDLINE: ten years on. *J Med Libr Assoc* 2006;**94**:130–6.

37. **Camosso-Stefinovic J.** *Developing Objectively Derived Search Strategies To Identify Randomized Controlled Trials in MEDLINE*: MSc thesis. Loughborough: Loughborough University, Department of Information Science, 2002.
38. **Bachmann L**, Coray R, Estermann P, *et al.* Identifying diagnostic studies in MEDLINE: reducing the number needed to read. *J Am Med Inform Assoc* 2002;**9**:653–8.
39. **Bachmann L**, Estermann P, Kronenberg C, *et al.* Identifying diagnostic accuracy studies in EMBASE. *J Med Libr Assoc* 2003;**91**:341–6.
40. **Wentz R**, Roberts I, Bunn F, *et al.* Identifying controlled evaluation studies of road safety interventions. Searching for needles in a haystack. *J Safety Res* 2001;**32**:267–76.
41. **Goss C**, Lowenstein S, Roberts I, *et al.* Identifying controlled studies of alcohol-impaired driving prevention: designing an effective search strategy. *J Inform Sci* 2007;**33**:151–62.
42. **Sladek RM**, Tieman J, Currow DC. Improving search filter development: a study of palliative care literature. *BMC Med Inform Decis Mak* 2007;**7**:18.
43. **The Concordance Software.** <http://www.concordancesoftware.co.uk> (accessed 5 Aug 2011).
44. **Buderer NMF.** Statistical methodology: I. incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Acad Emerg Med* 1996;**3**:895–900.
45. **Ingui BJ**, Rogers MAM. Searching for clinical prediction rules in medline. *J Am Med Inform Assoc* 2001;**8**:391–7.
46. **Matthews EJ**, Edwards AG, Barker J, *et al.* Efficient literature searching in diffuse topics: lessons from a systematic review of research on communicating risk to patients in primary care. *Health Libr Rev* 1999;**16**:112–20.