J Occup Rehabil (2009) 19:223–230
DOI 10.1007/s10926-009-9177-0

Precision and Recall of Search Strategies for Identifying Studies on Return-To-Work in Medline

Jean-François Gehanno · Laetitia Rollin · Tony Le Jean · Alexandre Louvel · Stefan Darmoni · William Shaw

Published online: 21 April 2009

Abstract *Introduction* The purpose of this study was to report on the qualities of various search strategies and keywords to find return to work (RTW) studies in the Medline bibliographic database. *Methods* We searched Medline for articles on RTW published in 2003, using multiple search strings, and hand searched 16 major periodicals of rehabilitation or occupational medicine. Among the retrieved articles, those considered to be relevant, were pooled in a Gold Standard Database. From this database, we identified candidate text words or MeSH terms for search strategies using a word frequency analysis of the abstracts and a MEDLINE categorization algorithm. According to the frequency of identified terms, searches were run for each term independently and in combination. We computed Recall, Precision, and number needed to read (NNR = 1/Precision) of each keyword or combination of keywords. *Results* Among the 8,073 articles examined, 314 (3.9%) were considered relevant and included in the Gold Standard Database. The search strings ("Rehabilitation, Vocational" [MeSH]), ("Return to work"[All]) and ("Back to work"[All]) had Recall/Precision ratio of 30.46/19.11, 59.55/87.38 and 3.18/90.91%, respectively. Their combination with the Boolean operator OR yielded to a Recall/Precision ratio of 73.89/58.44% and a NNR of 1.7. For the end user requiring comprehensive literature search, the best string was ("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH] OR "rehabilitation"[Subheading]), with a Recall of 88.22% and a NNR of 18. *Conclusions* No single MeSH term is available to help the physician to identify relevant studies on RTW in Medline. Locating these types of studies requires the use of various MeSH and non-MeSH terms in combination to obtain a satisfactory Recall. Nevertheless, enhancing the Recall of search strategies may lead to lower Precision, and higher NNR, although with a non linear trend. This factor must be taken into consideration by the end user in order to improve the cost-effectiveness ratio of the search in Medline.

Keywords Return to work · Medline · Bibliometrics

J.-F. Gehanno (✉) · L. Rollin · A. Louvel
Institute of Occupational Health, Rouen University Hospital, 1 rue de Germont, 76000 Rouen, France
e-mail: jean-francois.gehanno@chu-rouen.fr

T. Le Jean
Department of Rehabilitation Medicine, General Hospital of Dieppe, Dieppe, France

J.-F. Gehanno · S. Darmoni
GCSIS Laboratory, Rouen University Hospital, Rouen, France

W. Shaw
Liberty Mutual Research Institute for Safety, Hopkinton, MA, USA

Evidence based medicine (EBM) has become a landmark of medical practice for all medical practitioners. One of the pillars of EBM is to conduct effective literature searches. Finding useful information from a literature search involves two challenges. The first challenge relates to the amount of available information. In 1901, the *Journal of the American Medical Association* stated that "it is practically out of the question to be in touch with all the literature issued in any one department of medicine, because of the expense and time it would involve to attempt to master the stuff sufficiently well to separate the wheat from the chaff" [1]. This phenomenon has not waned since then and it was estimated in 1991 that the doubling time of the biomedical knowledge base was about 19 years [2]. The second challenge relates to

the complexity of the information to be retrieved and the time available for retrieval. It has been estimated that 7,287 articles relevant to primary care physicians were published monthly and that physicians would take an estimated 627.5 h per month to evaluate all these articles [3]. This is to be compared to the observation that internists declare spending 4.4 h per week reading medical journal articles and report reading only the abstract for 63% of the article [4].

Information retrieval is usually achieved by searching bibliographic databases. Of biomedical bibliographic databases, MEDLINE is the largest and most widely used in the world by physicians. This database is freely accessible through the Internet and is comprehensive, although not exhaustive, especially in the field of occupational health [5, 6]. Furthermore, it provides many tools, including a restricted thesaurus of medical titles, known as medical subject heading (MeSH) terms, which help users to recover the relevant information [7]. Nevertheless, identifying such information may be complex. A search should yield as much information as is available on a specific topic and lead to as few articles as possible that are unrelated to the search topic. In epidemiological terms this would mean a search that is both sensitive and specific, and in bibliometric terms this would mean a search that has both good Recall and high Precision ratios. A number of researchers have developed general strategies for MEDLINE searches of randomized controlled trials [8–11]; diagnostic studies [12, 13]; etiological, therapeutic, or prognosis studies [14], systematic reviews [15], or intervention studies in occupational health [16]. These strategies rely on the use of subject headings or text words that define the methodologies, clinical applications, and publication types in addition to the specific subject terms to achieve comprehensive yet accurate retrieval.

Disability prevention and return to work (RTW) is a daily preoccupation for many clinicians, including occupational physicians and physicians in rehabilitation medicine. Unfortunately, identifying studies on that topic in Medline is difficult due to the lack of appropriate keywords [17]. To date, no study has been performed to develop and compare literature search strategies that are both sensitive and precise to locate RTW studies. Thus, the purpose of this article is to report on the qualities of various search strategies and keywords to identify RTW studies in the Medline bibliographic database.

Methods

Construction of the Gold Standard Database

A four-step strategy was used to build the Gold Standard Database. The first step was to define what types of articles were to be included. We decided to include all the articles published in 2003 that had as an outcome the measurement of the rate of return to usual work of patients, or that assessed factors that impacted RTW. We excluded: (a) articles in which the return to usual activities, but not specifically work, was measured; (b) studies concerning only the effect of work or employment on a disease; and (c) descriptive studies providing only a rate of employment for a specific disease or group of people.

In a second step, we searched PubMed for articles that could fulfil our criteria. According to the MeSH database, the most fine-grained MeSH keyword intended to identify these articles was "Rehabilitation, Vocational", defined by the National Library of Medicine as "training of the mentally or physically disabled in work skills so they may be returned to regular employment utilizing these skills". Since we suspected that this MeSH keyword was not sufficiently sensitive, and in order to obtain the best possible Recall, we used two additional search strings: "Return to work"[All], and "Back to work"[All]. We identified additional relevant MeSH terms or subheadings in the articles that were retrieved with the first three search strings: "Sick Leave" [Mesh], "Rehabilitation" [Subheading], "Occupational therapy"[All], "gainfully employed"[All], "work status"[All], "RTW"[All], "Work re-entry"[All], "Employment outcome"[All] and "work retention"[All]. We therefore performed a new search combining all these words using the Boolean operator OR, and we limited our search to only those papers published in 2003 (2003[dp]). Overall, these search strings retrieved 5,434 articles. Two independent examiners examined all of these article listings (titles and abstracts) to select those relevant to our purpose. Discrepancies between search results were resolved by discussion and compromise.

The articles selected were then analyzed to identify the journals that contributed the most in the RTW field. This step was performed using MEDLINE Evaluator (MEVA), a medico-scientific data mining web service developed by the Institute for Medical Statistics and Epidemiology of the Technical University of Munich (http://www.med-ai.com/meva/index.html). In a third step, we hand searched the tables of contents and full volumes of several categories of biomedical journals published in 2003 for articles that fulfilled our inclusion criteria. The first category concerned the journals that had published at least 5 articles, identified at the end of step 2. These journals were the following: *Spine*, *Work*, *Journal of Occupational Rehabilitation*, *Surgical Endoscopy*, *Archives of Physical Medicine and Rehabilitation, Journal of Hand Surgery, Journal of Head Trauma Rehabilitation,* and *Occupational and Environmental Medicine*. The second category concerned the 4 major journals in the Rehabilitation category, according to the Journal of Citation Report 2003, namely *Physical Therapy*, *Supportive Care in Cancer*, *Journal of Electromyography and*

Kinesiology and *Archives of Physical Medicine and Rehabilitation*. The third category concerned the most important journals in the field of occupational health that were reported as containing the most information on occupational health problems [18]. This led us to five additional journals: *Contact Dermatitis*, *International Archives of Occupational and Environmental Health,* the *Journal of Occupational and Environmental Medicine*, the *Scandinavian Journal of Work, Environment & Health* and the *American Journal of Industrial Medicine*. Step 3 was performed independently by one of the authors. All but 3 of the relevant articles had already been identified in step 2. Altogether, we hand searched 16 different journals, and a total of 3,069 original research articles were reviewed from publication year 2003. In the final step, we merged all relevant documents retrieved to construct a Gold Standard Database.

Calculation of Recall, Precision, and Number Needed to Read (NNR)

The outcome of a search strategy is defined in the same terms as any diagnostic test (Table 1).

Recall and Precision are two accepted measurements to determine the utility of an information retrieval system or search strategy [19]. Precision can be seen as a measure of exactness or fidelity. It is defined as the number of relevant documents retrieved by a search divided by the total number of documents retrieved by that search. Recall is a measure of completeness, and is similar to the epidemiological concept of sensitivity. A sensitive search will find a high proportion of articles that belong to the Gold Standard Database resulting in the smallest number of false negatives, named Silence in bibliometric sciences. A precise search will exclude the highest proportion of articles that do not belong to the Gold Standard Database and yield the smallest number of false positives, named Noise in bibliometric sciences. Recall and Precision are inversely related which means they always have to be considered in combination.

We used two different approaches to identify the keywords to test. We analyzed the Gold Standard Database with a MEDLINE categorization algorithm to extract the most frequent MeSH terms used to index the articles included in our database [20]. Then, we analyzed the title and abstracts of the 314 articles of the Gold Standard Database with a Keyword Extractor (http://www.analogx.com/contents/download/network/keyex.htm) to identify the words that appeared most frequently. We selected the words that could be relevant for a Medline search.

We searched Medline using all the selected search words (MeSH and non MeSH) alone and then computed the Recall and Precision for each search strategy. Next, we built search strings with the combination that yielded the most sensitive search by subsequently adding those search words (with the OR operator) that found the most additional articles from the Gold Standard Database without decreasing Precision appreciably. A similar strategy was followed for Precision, by starting with search words with a Precision of at least 15% and Recall of more than 20% and adding search words (with the OR operator) that increased Recall with the least lowering of Precision. The number needed to read (NNR = 1/Precision) of each keyword or combinations of keywords was then calculated. We coined the term NNR in analogy to the number needed to treat (NNT) to describe the number of irrelevant references that has to be screened in order to find one of relevance [21]. The analyzes were performed with Excel (Microsoft).

Table 1 Method for calculating sensitivity, specificity, precision and recall of PubMed search strings for RTW studies

	Gold standard	
	+	−
Search term(s)		
+	a (true positives)	b (false positives)
−	c (false negatives)	d (true negatives)

Recall = sensitivity = a/(a + c); specificity = d/(b + d); precision = a/(a + b)

Results

The total number of articles retrieved by means of Medline and hand searching was 8,073. Overall, 314 (3.9%) articles were considered relevant and included in the Gold Standard.

The initial search strings ("Rehabilitation, Vocational" [MeSH]), ("Return to work"[All]) and ("Back to work" [All]) identified respectively, 197, 214 and 11 articles. When these keywords were combined using the Boolean operator OR, 397 articles were retrieved. Among them, 257 were considered relevant, according to our criteria. There was very little overlap between these keywords, as shown in Fig. 1.

The best single terms alone or in combination, either in terms of Recall or in terms of Precision, are displayed in Tables 2 and 3. The text word "Work", and its truncation ("Work*") allowed a Recall of more than 90%, but with a Precision lower than 1.5%. Higher Recall was attained with several combinations including the truncation of the text word "Work", but at the price of a Precision lower than 1%. Using the search string "Return to work", and searching into all search fields of PubMed, to retrieve articles concerning RTW allowed to identify 59.55% of relevant articles, with a good Precision (87.38%). Nevertheless, this search string missed 40% of relevant studies.

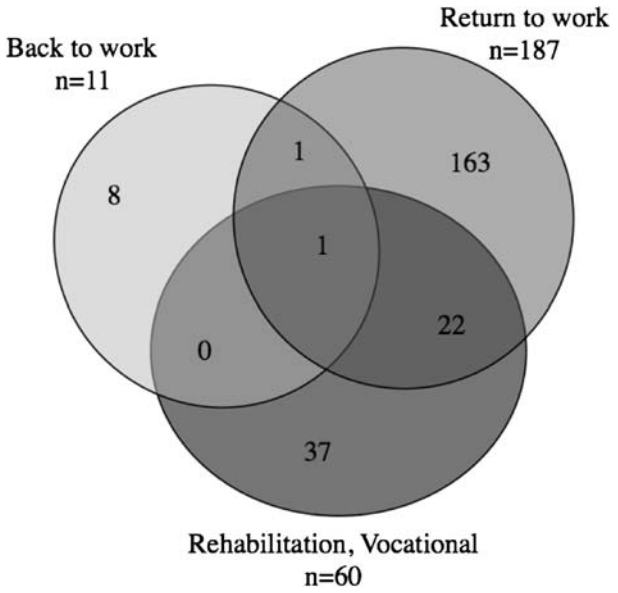

Fig. 1 Overlap between the keywords "Rehabilitation, vocational"[MeSH], "Return to work"[All] and "Back to work"[All] to retrieve relevant articles concerning return to work

For search strings, the highest Precision (90.91%) was obtained with the phrase "back to work", but using only this search string missed more than 94% of relevant papers. The best Precision when keeping Recall to 50% or more was observed with the strategy ("Return to work"[All] OR "Back to work"[All] OR "Rehabilitation, vocational"[MeSH]).

Table 4 shows the NNR for the strings leading to a Recall above 50%. The NNR vary between 1.1 and 289.9 which means that between 1 and 290 papers must be read in order to find a relevant article. Most often, a higher Recall implied a higher NNR. Nevertheless, it was not an absolute rule. For example, the string ("Return to work"[All] OR "Rehabilitation, vocational"[MeSH] OR vocational OR "Sick Leave"[MeSH] OR "work status"[All]) had a higher Recall that the string ("Return to work"[All] OR vocational OR "Rehabilitation, vocational"[MeSH] OR return), but a lower NNR.

Discussion

Return to work is considered of paramount importance as a quality-of-life outcome measure for many conditions, but also for societal concerns related to reduced productivity and cost of disability benefits. Therefore, the factors that could influence the rate of RTW for specific conditions or specific patients are increasingly taken into consideration, by rehabilitation or occupational physicians, but also by primary care physicians, under the request of patients or employers. Unfortunately, the literature in this field is difficult to identify [17].

No single MeSH term is available to help the physician, or the patient, to identify relevant information on that topic in Medline. The MeSH term proposed by the MeSH browser concerning RTW is "Rehabilitation, vocational". According to our results, its Recall is only 30.46%. The search strings ("return to work") and ("back to work") allowed Recall of respectively, 59.55 and 3.18, with good Precision. Unfortunately, even with the combination of these three strings, more than 25% of the relevant literature was missed. There is therefore a need for more exhaustive search strings. Our study documents those search strings that provide the best Recall, Precision, and best compromise between Recall and Precision, for retrieving articles concerning RTW from Medline.

One of the weaknesses of our study is that we searched only in the Medline database. It has been demonstrated than Medline is not exhaustive, especially in the field of occupational health [6]. Nevertheless, it is one of the most comprehensive bibliographic databases in the biomedical field, freely accessible through PubMed, which is not the case for Embase or Psycinfo. Furthermore, it can be used for some common questions in the practice of occupational medicine, for evidence based occupational medicine purposes [22].

Construction of the Gold Standard Database in this study was based on a different method than that used by previous studies of a similar nature. As Verbeek et al. [16] pointed out, nearly all the studies performed so far in this field have relied on hand searching of a limited, and somewhat arbitrary, subset of journals, which cannot be considered to be representative of the content of Medline as a whole. We gradually constructed our Gold Standard Database with a "snowball method" by using a preliminary choice of keywords, based on our experience in Medline searching, and adding new search words found in the title or abstract of the articles we identified as relevant. Since we also performed hand searching in the main journals in the field of RTW and occupational medicine, we consider that we did not miss important articles and that our Gold Standard Database may be considered to be as exhaustive as possible.

We were hampered by the lack of a computer program that could easily calculate test characteristics for different combinations of search words. Although we tried to evaluate the qualities of the search strategies as systematically as possible, we had to rely on subjective choices of search words to test, instead of being able to test all different combinations of search words. However, the method used, i.e. using the most frequent MeSH terms or most frequent words found in the title or abstract of the articles identified as relevant, seemed sufficiently reliable. In 2001, Haafkens

Table 2 Best single terms and combinations, in terms of recall, to retrieve articles published in 2003 concerning the return to work

	Recall (%)	Precision (%)
Single terms		
Work*[All]	93.95	1.04
Work[All]	90.13	1.47
Return[All]	66.56	9.22
"Return to work"[All]	59.55	87.38
Outcome[All]	47.45	0.29
Rehabilitation[Subheading]	42.99	2.83
Employ*[All]	42.68	1.08
Pain[All]	42.68	0.74
"Outcome assessment (health care)"[MeSH]	34.71	0.34
"Rehabilitation, vocational"[MeSH]	30.46	19.11
Disability[All]	29.94	2.66
Vocational[All]	21.02	20.06
Workers[All]	19.75	1.17
"Disability evaluation"[All]	19.43	4.45
Back[All]	19.11	1.18
Combinations[a]		
(Work* OR return OR "Return to work" OR outcome OR "rehabilitation"[Subheading] OR employ*)	100.00	0.34
(Work* OR return OR "Return to work » OR outcome OR "rehabilitation"[Subheading])	99.68	0.38
(Work* OR return OR "Return to work" OR outcome)	99.04	0.39
(Work* OR employ*)	98.73	0.80
(Work* OR return)	94.59	0.99
("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH] OR "rehabilitation"[Subheading])	88.22	5.55
("Return to work"[All] OR vocational OR "Rehabilitation, vocational"[MeSH] OR return OR "disability evaluation"[MeSH])	82.80	7.26
("Return to work" OR "Rehabilitation, vocational"[MeSH] OR vocational OR "Sick Leave"[Mesh] OR "work status")	82.17	35.20
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH] OR return)	78.34	9.51
("Return to work" OR "Rehabilitation, vocational"[MeSH] OR vocational OR "Sick leave"[MeSH])	77.71	34.96
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH]OR "disability evaluation"[MeSH])	77.71	15.85
("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH])	73.89	58.44
(Return OR back) AND work*	72.93	25.05
("Return to work" OR "Rehabilitation, vocational"[MeSH] OR vocational)	72.61	42.38

[a] In the combinations, all keywords were searched in[All Fields], otherwise specified. [All Fields] includes all search fields except for place of publication and transliterated title

* Truncation

et al. [6] studied the terms that could be used in Medline and 5 other databases to find relevant information on 5 chronic diseases and work participation. They excluded the words "work", "employment" or "occupation" because they were considered insufficiently precise. Based on our results, we would agree that excluding the words "work", "employment" or "occupation" significantly lowers the Recall. In fact, since the main purpose of their study was to identify databases, the authors considered that they were not able to construct a gold standard, and thus did not calculate Precision or Recall of keywords. Therefore, we feel that we did not overlook any important combination that would have changed our results substantially.

We did not calculate specificity because we considered it was not possible to be completely exhaustive in the search for articles involved in RTW. Some authors have hand-searched through a variety of journals, hoping that this allowed them to obtain the true negative results for the search strings, and therefore to calculate specificity [15, 16]. Our results, that confirm previous findings [5, 18],

Table 3 Best single terms and combinations, in terms of precision, to retrieve articles published in 2003 concerning the return to work

	Recall (%)	Precision (%)
Single terms		
"Back to work"[All]	3.18	90.91
"Return to work"[All]	59.55	87.38
"Work status"[All]	7.32	51.11
Vocational[All]	21.02	20.06
"Rehabilitation, vocational"[MeSH]	30.46	19.11
"Sick leave"[Mesh]	11.15	18.52
Return[All]	66.56	9.22
Absenteeism[MeSH]	4.78	8.57
Absenteeism[All Fields]	6.05	8.12
"Disability evaluation"[MeSH]	17.83	5.25
"Disability evaluation"[All]	19.43	4.45
"Occupational therapy"[All]	4.78	3.75
"Rehabilitation"[Subheading]	42.99	2.83
Disability[All]	29.94	2.66
Job[All]	13.38	2.31
Work[All]	90.13	1.47
Combinations[a]		
("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH])	73.89	58.44
("Return to work" OR "Rehabilitation, vocational"[MeSH])	68.15	55.15
("Return to work" OR vocational)	71.34	43.66
("Return to work" OR "Rehabilitation, vocational"[MeSH] OR vocational)	72.61	42.38
("Return to work" OR "Rehabilitation, vocational"[MeSH]) OR vocational OR "Sick Leave"[Mesh] OR "work status"	82.17	35.20
("Return to work" OR "Rehabilitation, vocational"[MeSH] OR vocational OR "Sick leave"[Mesh])	77.71	34.96
((Return OR back) AND work*)	72.93	25.05
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH] OR "disability evaluation"[MeSH])	77.71	15.85
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH] OR return)	78.34	9.51
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH] OR return OR "disability evaluation"[MeSH])	82.80	7.26
("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH] OR "rehabilitation"[Subheading])	88.22	5.55

[a] In the combinations, all keywords were searched in [All Fields], otherwise specified. [All Fields] includes all search fields except for place of publication and transliterated title

* Truncation

demonstrate that relevant articles are published in a wide range of journals; therefore, screening a limited number of journals can not be considered to provide an exhaustive overview of the literature in the field of RTW, as well as in other fields of occupational medicine. Furthermore, because the prevalence of RTW literature is limited in most of the journals, the use of a bibliographic database such as Medline is critical for both researchers and practitioners.

Our study observes a lower Precision of keywords than in many previous studies. Nevertheless, most, if not all, of the studies done so far to assess sensibility and specificity of search strings or keywords used a selection of journals to be hand searched. The Precision depends on the prevalence of the studies to be found. Therefore, the Precision will be much lower when searching the entire Medline, even for a specific year, as compared to searching in a range of journals selected because they publish more studies in the specific field studied than the average. By contrast to other studies on Medline [23, 24], the use of text words resulted in enhanced Precision and Recall, when compared to the use of MeSH terms.

The typical practitioner, who has limited time to access information from the literature, wants to find a substantial number of relevant articles without too much non-relevant information. Therefore, in an attempt to approach an optimal yield, it has been proposed to set the sensitivity

Table 4 Best single terms and combinations with a recall above 50%, in terms of number needed to read (NNR), to retrieve articles published in 2003 concerning the return to work

Single terms and combinations[a]	Recall (%)	Precision (%)	NNR
(Work* OR return OR "Return to work" OR outcome OR "rehabilitation"[Subheading] OR employ*)	100.00	0.34	289.9
(Work* OR return OR "Return to work" OR outcome OR "rehabilitation"[Subheading])	99.68	0.38	260.9
(Work* OR return OR "Return to work" OR outcome)	99.04	0.39	253.9
(Work* OR employ*)	98.73	0.80	124.3
(Work* OR return)	94.59	0.99	101.4
Work* [All]	93.95	1.04	95.9
Work [All]	90.13	1.47	68.1
("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH] OR "rehabilitation"[Subheading])	88.22	5.55	18.0
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH] OR return OR "disability evaluation"[MeSH])	82.80	7.26	13.8
Return[All]	66.56	9.22	10.9
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH] OR return)	78.34	9.51	10.5
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH] OR "disability evaluation"[MeSH])	77.71	15.85	6.3
((Return OR back) AND work*)	72.93	25.05	4.0
("Return to work" OR "Rehabilitation, vocational"[MeSH] OR vocational OR "Sick Leave"[Mesh])	77.71	34.96	2.9
("Return to work" OR "Rehabilitation, vocational"[MeSH] OR vocational OR "Sick Leave"[Mesh] OR work status")	82.17	35.20	2.8
("Return to work" OR vocational OR "Rehabilitation, vocational"[MeSH])	72.61	42.38	2.4
("Return to work" OR vocational)	71.34	43.66	2.3
("Return to work" OR "Rehabilitation, vocational"[MeSH])	68.15	55.15	1.8
("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH])	73.89	58.44	1.7
"Return to work"[All]	59.55	87.38	1.1

[a] In the combinations, all keywords were searched in [All Fields], otherwise specified. [All Fields] includes all search fields except for place of publication and transliterated title

* Truncation

threshold at a minimum of 65%, with a minimum Precision of 20% (NNR $\leq$ 5) [25]. In our study, the best compromise between Recall and Precision was provided by the combination ("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH]), with a Recall of 73.89% and a Precision of 58.44% (NNR 1.7). For experts, or for those interested in comprehensive retrievals or in searching for clinical topics with few citations, strategies with higher sensitivity may be more appropriate. To limit the number of articles to screen to some extent, it has been proposed to set the minimum Precision at 5% (NNR $\leq$ 20), under the condition that the sensitivity is at least 90% [16, 25]. Therefore, the best string for such end users appears to be ("Return to work" OR "Back to work" OR "Rehabilitation, vocational"[MeSH] OR "rehabilitation"[Subheading]), with a Recall of 88.22% and a NNR of 18. Usually, strategies with higher sensitivity lead to a lower Precision, and higher NNR. Nevertheless, this trend was not linear in our study (Table 4). For example, the strings ("Return to work"[All] OR vocational[All] OR "Rehabilitation, vocational"[MeSH] OR return[All]) and ("Return to work"[All] OR "Rehabilitation, vocational"[MeSH] OR vocational[All] OR "Sick Leave"[Mesh] OR "work status"[All]) had a Recall/NNR couple of 78.34%/10.5, and 82.17%/2.8, respectively. Using these two sensitive search strings we may expect to find 78 or 82% of all relevant material recorded on MEDLINE at a price of having to sift through 10.5 or 2.8 titles and/or abstracts to find one that refers to a relevant article on RTW. This difference appears to be acceptable until one realizes that, to find 50 relevant articles in MEDLINE, one will deal with 525 abstracts in the first case and "only" 140 abstracts in the second one, although the first option has a better Recall. This must be taken into consideration by the ends user in order to improve the cost-effectiveness ratio of their search.

Conclusion

Finding relevant studies on RTW in Medline requires the use of various MeSH and non-MeSH terms in combination to obtain a satisfying Recall, without sacrificing Precision.

One way to address this problem could be to create a new MeSH term. However, the number of MeSH terms cannot be expanded endlessly to satisfy any possible customer need. Text-word searching provides greater versatility and adaptability to particular research needs. Maximizing the potential of text-word searching requires titles and abstracts of the searched articles to be more descriptive and to use terminology more consistently. Authors of scientific publications in the field of RTW should be aware that the quality of titles and abstracts is likely to affect retrieval of their works from electronic databases, and could adopt the specific term "return to work" in a more consistent manner.

Acknowledgments Acknowledgment of financial support: this work was supported by Canadian Institutes of Health Research (CIHR) grant(s) FRN: 53909. The authors thank Richard Medeiros, Rouen University Medical Editor, for editing the final version of the manuscript.

References

1. The multiplication of books. JAMA. 1901;XXXVII:647–8.
2. Wyatt J. Uses and sources of medical knowledge. Lancet. 1991; 338:1368–72.
3. Alper BS, Hand JA, Elliott SG, Kinkade S, Hauan MJ, Onion DK, et al. How much effort is needed to keep up with the literature relevant for primary care? J Med Libr Assoc. 2004;92:429–37.
4. Saint S, Christakis DA, Saha S, Elmore JG, Welsh DE, Baker P, et al. Journal reading habits of internists. J Gen Intern Med. 2000;15:881–4.
5. Gehanno JF, Paris C, Thirion B, Caillard JF. Assessment of bibliographic databases performance in information retrieval for occupational and environmental toxicology. Occup Environ Med. 1998;55:562–6.
6. Haafkens J, Moerman C, Schuring M, van Dijk F. Searching bibliographic databases for literature on chronic disease and work participation. Occup Med. 2006;56:39–45.
7. Greenhalgh T. How to read a paper: the Medline database. BMJ. 1997;315:180–3.
8. Adams CE, Power A, Frederick K, Lefebre C. An investigation of the adequacy of MEDLINE searches for randomized controlled trials (RCTs) of the effects of mental health care. Psychol Med. 1994;24:741–8.
9. Nwosu CR, Khan KS, Chien PF. A two-term MEDLINE search strategy for identifying randomized trials in obstetrics and gynecology. Obstet Gynecol. 1998;91:618–22.
10. Watson RJ, Richardson PH. Identifying randomized controlled trials of cognitive therapy for depression: comparing the efficiency of Embase, MEDLINE and PsycINFO bibliographic databases. Br J Med Psychol. 1999;72:535–42.
11. Marson AG, Chadwick DW. How easy are randomized controlled trials in epilepsy to find on Medline? The sensitivity and precision of two Medline searches. Epilepsia. 1996;37:377–80.
12. Devillé WL, Bezemer PD, Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. J Clin Epidemiol. 2000;53:65–9.
13. Van der Weijden T, Ijzermans CJ, Dinant GJ, van Duijn NP, de Vet R, Buntinx F. Identifying relevant diagnostic studies in MEDLINE. The diagnostic value of the erythrocyte sedimentation rate (ESR) and dipstick as an example. Fam Pract. 1997;14: 204–8.
14. Felber SH. Searching for evidence-based oncology: tips and tools for finding evidence in the medical literature. Cancer Control. 2000;7:469–75.
15. Montori VM, Wilczynski NL, Morgan D, Haynes RB. Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. BMJ. 2005;330:68–71.
16. Verbeek J, Salmi J, Pasternack I, Jauhiainen M, Laamanen I, Schaafsma F, et al. A search strategy for occupational health intervention studies. Occup Environ Med. 2005;62:682–7.
17. Shaw W, Hong QN, Pransky G, Loisel P. A literature review describing the role of return-to-work coordinators in trial programs and interventions designed to prevent workplace disability. J Occup Rehabil. 2008;18:2–15.
18. Gehanno JF, Thirion B. How to select publications on occupational health: the usefulness of Medline and the impact factor. Occup Environ Med. 2000;57:706–9.
19. Dong P, Wong LL, Ng S, Loh M, Mondry A. Quantitative evaluation of recall and precision of CAT Crawler, a search engine specialized on retrieval of critically appraised topics. BMC Med Inform Decis Mak. 2004;4:21.
20. Darmoni SJ, Neveol A, Renard JM, Gehanno JF, Soualmia LF, Dahamna B, et al. A MEDLINE categorization algorithm. BMC Med Inform Decis Mak. 2006;6:7.
21. Bachmann LM, Coray R, Estermann P, Ter Riet G. Identifying diagnostic studies in MEDLINE: reducing the number needed to read. J Am Med Inform Assoc. 2002;9:653–8.
22. Verbeek JH, van Dijk FJ, Malmivaara A, Hulshof CT, Räsänen K, Kankaanpää EE, et al. Evidence-based medicine for occupational health. Scand J Work Environ Health. 2002;28:197–204.
23. Jenuwine ES, Floyd JA. Comparison of medical subject headings and text-word searches in MEDLINE to retrieve studies on sleep in healthy individuals. J Med Libr Assoc. 2004;92:349–53.
24. Chang AA, Heskett KM, Davidson TM. Searching the literature using medical subject headings versus text word with PubMed. Laryngoscope. 2006;116:336–40.
25. Schaafsma F, Hulshof C, Verbeek J, Bos J, Dyserinck H, van Dijk F. Developing search strategies in Medline on the occupational origin of diseases. Am J Ind Med. 2006;49:127–37.