A bibliometric study of the trend in articles related to epidemiology published in occupational health journals

Ken Takahashi, Tsutomu Hoshuyama, Kazuko Ikehami, Takashi Itoh, Toshiaki Higashi, Toshiteru Okubo

Abstract

Objectives—To study the role assumed by epidemiology in occupational health literature and characterise its change over the years.

Methods—A bibliometric study was conducted with a MEDLINE search to evaluate 9024 articles published in eight representative occupational health journals for the period 1980–93. Articles related to epidemiology were found by descriptors indexed in the articles and their numbers and proportions among all published articles was tabulated for the study period.

Results—The proportion of source items indexed by epidemiology as a descriptor increased over threefold from 7.9% (42/534) in 1980 to over 25% after 1990. Epidemiology was indexed most frequently as a subheading associated with other descriptors—for example, occupational diseases—epidemiology. Cohort studies had the largest increase in the descriptors that indicated epidemiological type of studies during the period studied.

Conclusion—Epidemiology is assuming increasing importance in occupational health literature. This was shown quantitatively and qualitatively by the use of descriptors in the MEDLINE database.

Keywords: epidemiology; bibliometrics; occupational health; MEDLINE

Epidemiologically oriented thinking has infiltrated the field of occupational health ever since Percival Pott.1 Occupational health authors, in their quest to summarise the knowledge acquired in this field, have undoubtedly been affected by this trend. However, the extent and manner in which epidemiological studies have increased in occupational health literature has remained obscure.

The number of scientific articles have been growing at incredible speed, and good coverage of the literature requires a computerised database. Recent advances in bibliometric evaluation make use of the functions designed primarily for the retrieval of articles from databases. In manually indexed databases, documents are characterised with vocabulary controlled by the indexers. The systems with which documents are indexed can, in turn, be used to synthesise information or measure a certain aspect of the relevant literature, hence the term “bibliometrics.” Recently, such a method was used to rank the productivity of research institutes by the number of publications and was introduced as “a novel type of assessment” in a commentary in the journal Nature.2 Another example is the assessment of frequently considered topics in occupational health journals in an effort to evaluate mutual influence between occupational and general medicine.3 Both examples used the MEDLINE database.

MEDLINE, produced by the United States National Library of Medicine (NLM), can be defined as the computerised version of the Index Medicus.4 MEDLINE is indexed using MeSHs (Medical Subject Headings), NLMs controlled vocabulary of over 16 000 terms.5 MEDLINE facilitates numerous types of searches of the medical literature through its various functions, including the descriptors indexed on journal articles to characterise their contents. “Epidemiology” is one of the medical subject headings assigned by national library of medicine indexers to characterise the content of various articles.

In this study, MEDLINE was used to conduct a bibliometric evaluation of articles published in occupational health journals. The objective was to assess, both quantitatively and qualitatively, the role of epidemiology in occupational health literature and characterise how it has changed over the years of a defined period.

Materials and methods

We conducted an on line search of DIALOG (currently Knight-Ridder) MEDLINE, file 155, which contains records from 1966 to the present. MEDLINE records are indexed with MeSHs with an average of 10–12 descriptors an article.6 Descriptors are subject index words and phrases assigned to records to characterise the content of the original document.7

SELECTION OF MATERIALS

McCunney and Harzbecker showed that “occupational diseases” was the most frequent topic in articles published by six occupational health journals.8 With the RANK DESCRIPTOR function in MEDLINE, we also confirmed that the most commonly indexed topic (or descriptor) was occupational diseases in these journals. Based on this finding, objective criteria for selecting occupational health journals were developed for this study as follows: (a) that the journal ranked among the top journals in the number of papers published on occupational diseases (RANK JOURNAL function); (b) that the
journal be categorised under “public health” in the SCI Journal Citation Reports; and (c) that the journal began publication in 1980 or before. The application of these criteria resulted in the selection of the eight occupational health journals (table). There were three journals which ranked high for criterion (a) but were excluded for not satisfying (b) or (c). The final list of journals chosen for the present study included the six occupational health journals in the paper by McCunney and Harzbecher and were considered reasonably representative of occupational health literature.

Articles published in these journals (source items) were limited to articles written in English (language code; LA = ). Reviews, editorials, and letters for document type (DT = ) were excluded. This was thought necessary to focus as much as possible on papers containing the results of substantive research— that is, original research articles. As a result, a total of 9024 source items from eight occupational health journals for the period 1980–93 were identified for evaluation (table).

**IDENTIFICATION OF ARTICLES RELATED TO EPIEDEMILOGY**

In the MEDLINE search, the SELECT (or S) command was given in combination with a descriptor defined in MeSHs to identify journal articles indexed by that descriptor. MEDLINE records were printed out in full or part to confirm the indexed information as necessary.

The manoeuvre to retrieve articles indexed by descriptors is exemplified as follows:

(a) **SELECT**(OR S) OCCUPATIONAL DISEASES/DE

(b) **SELECT**(OR S) EPIDEMIOLOGY/DE

(c) **SELECT**(OR S) OCCUPATIONAL DISEASES (L)EPIDEMIOLOGY/DE

The /DE suffix limits the search to articles indexed by the respective descriptors. If the /DE suffix is omitted, the search will include scanning of title and abstract as well as the descriptor field for an exact match of word strings (compound terms such as occupational diseases will be searched as a descriptor regardless, if the /DE suffix is omitted). Manoeuvre (a) retrieves articles indexed by the descriptor occupational diseases and (b) retrieves articles indexed by the descriptor epidemiology. Both descriptors are defined MeSHs. However, epidemiology is most frequently indexed as a subheading— that is, a term attached to other descriptors to specify a certain aspect of a subject. Epidemiology is one of the 80 topical subheadings used currently in MEDLINE, other examples being drug therapy, economics, diagnosis, etc. Subheadings included in the descriptor field are searchable as individual terms or as bound phrases. Manoeuvre (c) retrieves articles indexed by the descriptor combining the main heading and subheading as in occupational diseases— epidemiology.

Descriptors representing the main emphasis of the article are called major descriptors and designated by an asterisk (*) in front of the descriptor. It is under these headings that the citation can be located in Index Medicus. The manoeuvre for retrieval was further limited to records indexed by a major descriptor by using the /MAJ suffix.
Results

The table shows the number of source items, and number and proportion of source items indexed by epidemiology both as a descriptor and a major descriptor published by the eight occupational health journals for the study period. Of the eight journals, Br J Ind Med was published quarterly until 1984 and monthly since then; it changed its title to Occup Environ Med in 1994. The J Soc Occup Med changed its title to Occup Med (Oxford) in 1992.

Of the 9024 papers published in the eight journals for the period 1980–93, 1597 (17.7%) were indexed by the MeSH “epidemiology” as a descriptor and 842 (9.3%) as a major descriptor. The annual trend in the proportion of source items indexed by epidemiology as a descriptor increased over threefold from 7.9% (42/534) in 1980 to over 25% after 1990. Similarly, as a major descriptor, its proportion increased almost twofold from 5.2% (28/534) in 1980 to 10.2% (81/792) in 1993 (table and fig 1). Among the eight occupational health journals, the Am J Ind Med published the largest number (371) and proportion (371/1325; 28.0%) of articles indexed by epidemiology. The Arch Environ Health (24.0%) and the Scan J Work Environ Health (23.1%) followed in terms of proportion. The annually increasing trend in this proportion was most evident for the Br J Ind Med.

The annotated list of MeSHs cites epidemiology both as a subheading and a term representing specialty, subheadings are cited as the preferred usage. The MEDLINE search confirmed that source articles related to epidemiology were, in most cases, indexed by MeSHs such as occupational diseases, neoplasms, and lung neoplasms and were bound with the term epidemiology as a subheading. Hence, articles were assigned descriptors in the form of bound terms such as occupational diseases—epidemiology, neoplasms—epidemiology, and lung neoplasms—epidemiology. A small fraction of source items retrieved by EPIEMIOLOGY/DE, however, consisted of recent articles indexed by newly defined descriptors containing the
term epidemiology, such as confounding factors (epidemiology) and bias (epidemiology).

Occupational diseases—epidemiology was the single most frequently indexed descriptor bound with epidemiology, and it had increased in proportion of source articles from 3-9% (21/534) in 1980 to 10-6% (84/792) in 1993. Similarly, neoplasms—epidemiology and lung neoplasms—epidemiology increased from 1-9% (10/534) to 5-8% (46/792) and 0-6% (3/534) to 2-5% (20/792), respectively, during this period (data not shown).

In figure 2, the proportion of source articles indexed by descriptors indicating an epidemiological type of study was evaluated for the same period. It should be noted that the denominator of the proportion evaluated here is the number of source articles indexed by the descriptor epidemiology and not the number of source items themselves. Descriptors indicating study type were grouped into three major categories based on the MeSHs tree structure, under study characteristics (not MeSHs) for epidemiological methods. That is, cohort studies, follow up studies, longitudinal studies, and prospective studies were grouped into cohort studies, case-control studies and retrospective studies into case-control studies, and cross sectional studies as they were. Overall, the proportion of source articles indexed by any of the descriptors indicating epidemiological type of study increased from 16-7% (7/42) in 1980 to 64-8% (136/210) in 1993. Among the major categories of study type, cohort studies showed the largest increase from 4-8% (2/42) in 1980 to 22-9% (48/210) in 1993, whereas case-control studies were indexed for 11-9% (5/42) in 1980, decreased during the middle of the study period, and increased to 14-3% (30/210) in 1993.

Discussion
A MEDLINE search of source items in representative occupational health journals showed that the proportion of source items indexed by epidemiology as a descriptor increased over threefold during the period 1980–93. Also, epidemiology was indexed as the main focus (major descriptor) twice as often in the later years than in the earlier years of the period investigated. It can thus be inferred that epidemiology is assuming increasing importance in occupational health literature. However, the proportional increase was greater for articles indexed by epidemiology as a descriptor than as a major descriptor. Hence, epidemiological methods are being introduced increasingly into occupational health studies without necessarily constituting the major focus of the study. This may be a reflection of the increase in the combined application of epidemiological methods and other research methods. There are several points which prompt examination in making these inferences.

Firstly, it is possible that the apparent trend reflects changes in the practice of indexing articles. Historically, the MeSH annotated alphabetic list6 started indexing epidemiology as a subheading in 1989, until which time "occurrence" had been indexed for the same purpose since 1966. However, MeSH descriptors are revised annually and all MEDLINE records are updated yearly with these changes.7 In effect, epidemiology was indexed as a subheading to all pertinent articles of the study period. Nevertheless, the rising trend from 1988 onward may have occurred for this reason.

To further explore this issue, source items in the general medical literature were examined by a compatible method to assess whether the proportion of articles indexed by epidemiology as a descriptor increased as well. This proportion increased from 2-1% in 1980 to 6-2% in 1993 with a noticeable increase around 1990, thus showing a similar annual trend to the use of epidemiology indexed as a subheading, but at a lower proportion and rate of increase. Thus, although the introduction of epidemiology as a subheading is likely to have influenced the indexing procedures at large, the trend in the occupational health literature is more distinct, and supports the notion that, over the years, the role of epidemiology in occupational health research has grown.

Second, the validity and reliability of the
A bibliometric study of the trend in articles related to epidemiology published in occupational health journals

A bibliometric study trend in The MeSHs epidemiology these were such epidemiological aspect have in descriptors indexers, items indexed in the database. Within the boundaries of this limitation, descriptors assigned to journal articles can be used to form an overview of the literature. Bibliometric studies should be investigated with experience, expertise, and caution. Nevertheless, the indexed information such as the descriptor epidemiology will eventually be used to retrieve the article from the database, which fact alone justifies the use of this method.

The MeSH epidemiology as a subheading is always coupled with another MeSH, so the trend in indexing of bound terms was also examined. The indexing rule is that the subheading epidemiology can be bound with MeSHs for diseases and geography only. Occupational diseases, neoplasms, and lung neoplasms were the three most common MeSHs indexed in combination with the subheading epidemiology. The increase in proportion of source items indexed by epidemiology reflects the combined effect of these descriptors. It is noteworthy that about 25% or more of the recent articles published in occupational health journals deal with the epidemiological aspect of occupational diseases.

The recent development of epidemiological methods have prompted investigators in general medicine, pediatrics, and psychiatry to examine the type of epidemiological study design used for studies in their respective journals. Such tasks were pursued by subjective review of a sample of articles from the literature. In this study, descriptors indicating an epidemiological type of study were examined objectively for their proportional trends (fig 2). The analysis was first restricted to source items indexed by epidemiology, then articles indexed by descriptors indicating epidemiological type of study were identified and grouped into three distinct groups. The advantage of this procedure over showing crude proportions is that source items indexed by epidemiology (=100%) could be divided into proportions. This was considered to outweigh the fact that articles not indexed by the descriptor epidemiology are precluded to avoid overlap counting for the same article caused by multiple term indexing.

A biased total proportion of articles indexed by any of the descriptors indicating epidemiological type of study showed an increase over the years. It is possible that a spreading recognition of epidemiology enhanced its characterisation in indexing procedures, or researchers are increasingly using definable epidemiological methods. We speculate that many factors of this nature interact in producing this trend.

Epidemiological study types such as case-control studies, cohort studies, and cross sectional studies are listed at the same level under analytical studies (epidemiology) (not MeSHs) subcategorical to study characteristics (not MeSHs) for epidemiological methods. The classification scheme to place the three types of study at the same level may be justifiable, or at least acceptable, from a standpoint of classification by direction—that is, prospective, retrospective and cross sectional studies—notwithstanding its nomenclature.

However, classifying retrospective studies under case-control studies and follow up studies, longitudinal studies and prospective studies under cohort studies is debatable, for the mere reason that retrospective studies are not necessarily case-control studies, follow up studies are not necessarily cohort studies, and so forth. This is pertinent when a MEDLINE search is performed in the direction of broad to narrower terms (such as the EXPLODE(!) feature). Standardisation in terminology of research methods on the part of researchers would also improve the quality of indexing, the efficiency of database retrieval, and hence the potentiality of bibliometric evaluation.

Cohort studies were found to have increased more than the other two study types in this investigation. In general, studies of a prospective nature (including cohort studies) fare better for their use as inferential research tools that cross sectional studies. Although cohort studies by strict epidemiological definition may not be exactly represented, a trend in occupational health research with an inclination towards stronger study design in exploring disease aetiology is suggested.

As the analysis used predefined indices available in the database, a conventional, more detailed review is needed to characterise the epidemiological aspect of articles related to epidemiology. However, the option for a conventional type of review is at the expense of unbiased judgment. Finally, although inferences were confined within the rules on indexing procedures and MeSHs terminology, the use of descriptors in a MEDLINE search provided a unique time and cost effective opportunity to obtain an unbiased overview of the trend in articles related to epidemiology in occupational health literature.

Vancouver style

All manuscripts submitted to *Occurop Environ Med* should conform to the uniform requirements for manuscripts submitted to biomedical journals (known as the Vancouver style.)

*Occurop Environ Med*, together with many other international biomedical journals, has agreed to accept articles prepared in accordance with the Vancouver style. The style (described in full in the *BMJ*, 24 February 1979, p 532) is intended to standardise requirements for authors.

References should be numbered consecutively in the order in which they are first mentioned in the text by Arabic numerals above the line on each occasion the reference is cited (Manson confirmed other reports). In future references to papers submitted to *Occurop Environ Med* should include: the names of all authors if there are seven or less or, if there are more, the first six followed by et al; the title of journal articles or book chapters; the titles of journals abbreviated according to the style of Index Medicus; and the first and final page numbers of the article or chapter. Titles not in Index Medicus should be given in full.

Examples of common forms of references are:


References should be numbered consecutively in the order in which they are first mentioned in the text by Arabic numerals above the line on each occasion the reference is cited (Manson confirmed other reports). In future references to papers submitted to *Occurop Environ Med* should include: the names of all authors if there are seven or less or, if there are more, the first six followed by et al; the title of journal articles or book chapters; the titles of journals abbreviated according to the style of Index Medicus; and the first and final page numbers of the article or chapter. Titles not in Index Medicus should be given in full.

Examples of common forms of references are: