

Medical indexing in the United States

Janyne Ste Marie

This article discusses the US market for medical indexing, education, training, common issues in indexing medical topics, tools used in the indexing process, and indexing standards.

Introduction

Medical indexers serve the portion of the publishing industry related, of course, to medicine. At the time of writing there were approximately 168 medical publishers in the United States (www.literarymarketplace.com), which together publish thousands of titles each year. The subject matter covers many medical specialties, from acupuncture, dentistry, and clinical and research medicine to forensic pathology, neurology, pediatrics, psychiatry, and all points in between. Indexers who specialize in medicine must have a good science background and an ability to extract relevant information from a given passage. They must also have access to a variety of dictionaries and other reference material. Medical indexers work in a variety of formats, providing indexes for books, periodicals, abstracts, CD-ROMs, audio recordings, graphics, and product catalogs. According to the *ASI Salary survey* of 2004, 14 percent of the indexers in the United States report medicine as a specialty (ASI, 2004).

Education and training

According to ASI (2004), 33 percent of indexers are self-taught, 43 percent report employer-sponsored training, and 47 percent trained via the USDA Basic Indexing course (a self-taken correspondence course provided by the US Department of Agriculture). Many indexers received training through more than one avenue, and reported this to the survey. In addition, 48 percent reported training via mentoring relationships. Other sources of training were library school curricula, apprenticeships, video courses, and various other indexing courses taught by individual indexers.

Many indexers in the field of medicine have practical experience as nurses, laboratory or hospital personnel, or as physicians in various specialties. Others have a formal university education that includes medical coursework or a background in another field of science, such as biochemistry or engineering. Continuing education is strongly encouraged and available through many sources, most notably the ASI annual conference and workshops organized by the various indexing societies. It is also considered good practice to keep up with publications in specific fields of interest, just as other medical professionals do.

Common issues in medical indexing

As with indexing in any field, term consistency is a constant issue, most notably in works with multiple authors. As an

example, a few years ago I indexed a book entitled *Vascular dementia*. Among the many types of dementia referred to were vascular ischemic dementia (VID) and ischemic vascular dementia (IVD). A query of *Taber's dictionary* (Taber, 1997) showed the two to be one and the same, while a query of the PubMed online database (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>) showed that the preponderance of research used the IVD variation. Upon presenting this dilemma to my production editor, I was instructed to stick with each individual author's usage. This left me with headings for both variations, with cross-references from each to the other.

This example also serves to highlight the role played by the client in index creation. As a professional indexer I follow the Golden Rule: the publishing house with the Gold makes the Rules. The chosen resolution was not my first choice, but it was my client's preference. Other examples of the term consistency issue include references to Parkinson's disease/Parkinsonism, heart attack/myocardial infarction/acute ischemic cardiac event, and cancer/neoplasia/tumor. Medical procedures, causal organisms, and nearly every portion of the human anatomy can also be referred to by different terms.

Since term consistency is highly desirable in most indexes, this problem must be recognized by the indexer and resolved. Usually I choose the more common usage, with cross-references from the others. For example, in a book that discusses tumors, neoplasia, and cancer, I would first index each term at its occurrence in the text. At the end of the reading phase of indexing, I would determine which heading has the largest entry, consider the audience for the book, then choose which term is most appropriate. Natural language (a term such as cancer or tumor) would be my first choice in a book for lay readers, whereas I would choose the technical variant (such as cancer or neoplasia) for research scientists. But of course, as I have said, the client's preference is also a big factor in this decision.

Medicine is full of specialized terminology. There are scientific binomial names, chemical names, and names of genes, to give a few examples. Some standardization must be maintained in the alphabetization of these terms, or no one will know where to look for them. How do you file (1R-2R)-(-)-2-Amino-1-(4-nitrophenyl)-1,3-propanediol? Where do *c-Myc* and *v-Myc* entries go? There are many sources of helpful information in this regard. The SI occasional paper 3, 'Indexing the medical sciences' (Blake et al, 2002) is widely used in the United States. Equally useful are catalogs from companies that deal in medical products, such as Merck and Sigma-Aldrich,¹ as is MeSH, the Medical Subject Headings database (www.nlm.nih.gov/mesh/), maintained by the US National Institutes of Health.

In addition to filing order, proper formatting of entries can be quite problematic. The science of genetics is especially cumbersome; as an example, a gene name is commonly formatted in italics, as in the *Myc* examples above, or the human *SRC* gene. The protein product encoded by the gene, however, is commonly expressed in a roman or other nonitalic typeface as Src protein. The determination of which form an author is referring to in a given passage is one of medical indexing's more interesting challenges. Additionally, there are different conventions to consider with respect to species. The SI publication gives these guidelines:

- Human genes should be printed in uppercase italics: *BRCA1*, *CYP1A2I*.
- Mouse and rat genes should be printed in lower-case italics, with an initial capital letter: *Brca1*, *Cyp1A2*.
- Yeast genes should be printed in upper-case italics for dominant alleles, and lower-case italics for recessive alleles: *ARG2*, *arg2*.
- Bacterial and viral genes should be printed in lower-case italics, but may include upper-case letters: *lacZ*, *cagA*.

(Blake et al, 2002: 49)

Biochemistry also has term consistency issues. Some compounds, such as DNA, are normally indexed under the abbreviation. Other compounds, notably organic compounds, have different naming conventions which authors use interchangeably. Examples of this include citric acid/citrate, and formic acid/formate. These forms should not be indexed separately, but rather together, with cross-references given at the alternative forms where appropriate. I must again emphasize the need for the professional indexer to recognize and resolve these issues.

Tools of the medical trade can also be challenging. If you have a limited familiarity with surgical instruments, how can you know whether an Allis clamp is the same as an Allis forcep? Again, medical indexers must be vigilant and know when to consult references. In addition to bookshelf dictionaries, there are online discussion groups and even online catalogs from medical supply companies that provide pictures of many instruments. Merck manuals and Sigma-Aldrich catalogs, mentioned previously, are helpful here also.

Tools of the indexing trade

Indexing in the United States at least is production work; there is always a deadline and our clients want the best index that can be produced for the lowest cost in the (usually inadequate) amount of time allotted. Medical indexers are certainly not exempt from these pressures. With the advent of computers, our work has of course been revolutionized. Every medical indexer I know uses some form of dedicated indexing software, more often than not Cindex™, Macrex™ or Sky Index™. The facilities offered by these packages are a tremendous boon to medical indexers in dealing with the problems discussed above. I find myself, for example, constantly using both the group function in Cindex to isolate all records containing, say, 'IVD', and the Boolean operators

AND, OR, and ONLY to bring together terms in various combinations (for example, cancer OR tumor OR carcinoma OR neoplasm).

Many publishing houses now send PDF files to the indexer. The use of these files makes data entry a much easier task, allowing the indexer to copy and paste index entries. This is especially useful in the cases of cumbersome disease names, author indexes, and the names of medicines, genes, and other chemical substances. PDF files are also useful for searching for passing mentions of terms. Many publishers fail to provide the Table of Contents for the book; I've been faced many times with the prospect of searching for mentions of a topic that seemed quite inconsequential at first reading, but actually had an entire chapter devoted to it later in the book.

Standards

As has been stated by many, the wonderful thing about standards is that there are so many to choose from (Browne and Jerney, 2007). In the United States, the most commonly accepted standard is the *Chicago manual of style*, currently in its 15th edition (2003). ISO 999: 1996 is also a very common standard. Different publishing houses have their own internal style guides, of course. For medical indexing, the SI publication (Blake et al, 2002) is a common guideline; the American Medical Association also has a style guide, currently in its tenth edition.

Conclusion

Medical indexing is an interesting and challenging pursuit. Workers in the field have a wide variety of backgrounds and must solve many complex problems. We work in many different fields of medicine that are considered lifetime specialties for individual practitioners, and continuing education is a requirement for success. Although some researchers are attempting to construct indexing programs to replace human indexers, no one to date has succeeded. Until artificial intelligence with indexing capabilities becomes a reality, human indexers, even in these hard times, can rely on steady work in this field.

Acknowledgment

This is a shortened and updated version of an article previously published in *Information Wissenschaft&Praxis (IWP)*, 8-2007, reprinted here by kind permission of the editor.

Note

- 1 These are available from Sigma-Aldrich, www.sigma-aldrich.com and Merck, www.merck.com respectively. Merck manuals are not free of charge; Sigma catalogs are.

References

- American Medical Association (2007) *Manual of Style*, 10th edn. New York: Oxford University Press.

American Society of Indexers (2004) *ASI Salary Survey*. Wheat Ridge, Colorado: ASI.

Blake, D., Clarke, M., McCarthy, A., and Morrison, J. (2002) *Indexing the medical sciences*, Occasional Papers on Indexing no. 3. Sheffield: Society of Indexers.

Browne, G. and Jermy, J. (2007) *The indexing companion*. Melbourne: Cambridge University Press.

Chicago (2003) *Chicago manual of style*, 15th edn. Chicago: University of Chicago Press.

Merriam-Webster (nd) *Merriam-Webster Dictionary Online* [online] www.m-w.com (accessed May 2007).

Mulvany, N. (2005) *Indexing books*, 2nd edn. Chicago: University of Chicago Press.

Stauber, D. M. (2004) *Facing the text: content and analysis in book indexing*. Eugene, Ore.: Cedar Row Press.

Taber (1997) *Taber's Cyclopedic Medical Dictionary*, 19th edn. Philadelphia, Pa.: F. A. Davis.

Janyne Ste Marie is a freelance indexer and owner of the Electronic Scribe Indexing Service. She lives in Green Bay, Wisconsin, USA. She is a graduate of the University of Wisconsin, 1989 with a Bachelor of Science degree in biology, and obtained an indexing certificate from the USDA Graduate School in 2002. She has been indexing in the medical sciences since 2002. Email: jstemarie@new.rr.com

SKY Index™ Professional Edition

The fastest way to complete your index

SKY Index is easy to learn and includes data entry and editing features that allow you to enter and edit your index *faster*. Of course, the faster you complete your index, the greater your potential earnings. Don't take our word for it, though. Try SKY Index today! Download a **FREE** demo from our web site and find out why everyone is talking about SKY Index.

A Student Edition is also available. Call for details.

SKY
Software

www.sky-software.com
(540) 869-6581

sales@sky-software.com • SKY Software, 114 Lariat Court, Stephens City, VA 22655

Greatest editing flexibility

SKY Index has the most complete set of editing commands and tools. SKY Index's spreadsheet approach to data entry and editing provides editing capabilities simply not possible with other data entry techniques.

Easiest to learn and configure

Because of our spreadsheet approach, many indexers are able to effectively use SKY Index with virtually no learning curve. When configuring your index, preview screens are frequently provided so that you don't have to guess what effect your changes will have. Presets for commonly used settings are also provided.

Most advanced AutoComplete

SKY Index's AutoComplete feature uses sort order, frequency of use, and last time of use when anticipating your entries resulting in greater accuracy. AutoComplete also works for cross-references and it can even be used as a controlled vocabulary.

Drag and drop embedding

Simply drag your index entries from SKY Index and drop them into your Word document.

Already have indexing software?

We offer a discount to owners of popular indexing programs. Please visit: www.sky-software.com/upgrademe for more information.

Minimum System Requirements: 500 MHz Pentium; Windows 98 or higher; 32MB RAM; CD-ROM; word processor.
SKY Index is a trademark of SKY Software. SKY Software logo is a registered trademark of SKY Software. Windows is a trademark of Microsoft