COMPUTER SOFTWARE REVIEWS

The Merck Index Twelfth Edition on CD-ROM Version 12:1 1996 for Microsoft Windows[1]

Carmen Nitsche

Nalco Chemical Company, 1 Nalco Center, Naperville, Illinois 60563

Received June 30, 1997

To me the Merck Index has always seemed like a chemist's necessary companion. I received a copy from a good friend after I finished my B.A. in chemistry. Since then I have kept a copy of the Merck Index above my desk through graduate school and all of my professional positions. I have not, however, kept up with every new edition, and I was curious to see how much I would gain by upgrading to the electronic version of this resource.

The CD-ROM version reviewed here is produced jointly by Merck & Co. and Chapman and Hall. Like the book, it covers more than 10 000 monographs, each representing either a single substance or a family of related substances. The monographs include, as defined in the manual, human and veterinary drugs, biologicals, natural products, agricultural compounds, commercial and laboratory chemicals, and environmentally significant compounds.

But, unlike the book, the CD-ROM version allows you to search the contents in many new ways, including chemical structure and substructure. It is worth noting that, although the Merck Index is available electronically on various commercial online vendors, none allow direct substructure searching of the Merck Index compound collection. You can also range search property fields, such as melting point or boiling point, and further combine those results with a chemical name, a therapeutic category, or a substructure fragment. Results can then be reviewed on-screen in various formats (including one that looks just like the book record), printed off, or saved to disk.

It took me less than a minute to load the system, and all I had to read was a one-page install guide in the front of the manual that accompanies the disk, which is about as much effort as I wanted to make. The manual includes the following hardware and software requirements: an IBM compatible PC with Intel 80486DX, 8Mb RAM, a VGA monitor (preferably color), and 40Mb hard disc with 7Mb free and a CD-ROM drive, running MS-DOS 3.3 or higher, MSCDEX version 2.1 or higher, and Microsoft Windows version 3.1 or Windows 95. If you must print out the information, you also need a printer compatible with PostScript or Hewlett Packard configurations.

When it comes to searching, you are given four options: **Quick Search**, **Menu Search**, **Command Search**, and **Structure Search**. For a quick lookup of a chemical, **Quick Search** is generally adequate. By selecting this option you are lead to a screen that asks for what field you want to search, which include **All Names**, **Molecular Formula**, **Properties**, **Therapeutic Category**, **All Text**, **CAS Registry Number**, **Use**, and **Title Name**. Selecting the appropriate field takes you to a screen that includes an index of every term posted in that field. You can enter your term directly into the search box or start typing in the Index Stem box, which will allow you to browse through your term's near neighbors as well. Once you find the desired term just double click to get that term into the search box.

I discovered that you will probably want to check the stem box first. This will save time, since you do not always know what synonyms are present in the database. Entering simply butene into the search box left me with zero hits. Entering butene in the stem box gave me 1- and 2-butene as options to be searched. I also discovered that even if you search for trichloromethane, you will only find the monograph name "chloroform" on your results list.

If you like to take full control of your search, then you can use the **Command Search**. But if you want to combine multiple search concepts and you are not a reader of manuals, then the **Menu Search** screen will guide you through the Boolean operators. Although you can use **&** for "and", / for "or", and \ for "not" in all entry screens, the menu option takes care of them automatically.

Regardless of the text search mode you prefer, you will want to become familiar with the word truncation symbols, however, since they offer a powerful retrieval option. * can be used not only for right multicharacter truncation but also for left and internal multicharacter truncation. A single wild character within a word is denoted by a **?**. Depending on your system, be prepared to wait awhile for your results if you truncate or range search. On my albeit slow system an **All Name** search of ***chloro*** took 5 min. The same search ran in just over a minute on one of my group's faster systems. Some of my range searches of boiling point of melting point took long enough to cause me to abort the searches.

The fourth search option is by chemical structure. If you ever use STN-Express to search STN you will recognize the structure building interface. Drawing structures is fairly straightforward. You can run a structure search independently or in combination with a text search, a powerful capability that brings out a whole new way to use the Merck Index. To maximize the speed and accuracy of your structure search go through the Query Definition phase to define attributes such as embedded and/or isolated rings, exact or normalized bonds, and ring or chain nodes. Given the restricted size of the database, however, very often the defaults will work just fine. Processing times were reasonable: I ran a search for a five-carbon chain, which resulted in over 950 qualifying candidates in about 2 min. I would recommend reading the section in the manual on structure searching, especially on how to initiate a search and how to combine a structure search with text searching.

I have a few gripes with the software. There is no hit-term highlighting. There are also a lot of icons to contend with. Six pages of the 55 page manual list each screen and all the icons that go with that screen (although many of the icons appear in multiple screens). In the text search portion of the program help balloons appear as you move over the icons; while in the structure building part the icon definitions appear in the screen's lower bar. I also consider the combination structure/text search not as obvious as one might wish, and you will have to experiment a bit (and/or read the manual carefully) to utilize this feature properly. I never got the mark feature to work on my system, but in all fairness I did not pursue this with customer service.

My only serious concern with this product is unrelated to content or search capability but rather to price. I am comfortable spending $45 for a reference book that sits at my desk. At $250 for the CD-ROM I would almost be willing to walk to the library (if my office were not already there) when I needed a structure search. Although an online STN Registry File structure search is nearly $100 a pop, it encompasses a much larger database with much more current information.

This CD is not just a glorified version of the book, and I would expect every library to get a copy. As for getting a personal copy, ask yourself the following questions to determine if you need this resource at your desk:

1. Does the Merck Index cover the realm of compounds I am primarily interested in?

2. Do I need to identify compounds by combining multiple concepts, such as therapeutic category with melting point and some structural feature?

3. Will I need to carry out such complex searches routinely?

If you answer yes to the above questions then you should probably consider getting a desk copy of the Merck Index CD-ROM as well.

REFERENCES AND NOTES

(1) The Merck Index Twelfth Edition on CD-ROM Version 12:1 1996 for Microsoft Windows is available from the Electronic Publishing Division, Chapman & Hall Inc.: 115 Fifth Avenue, New York, NY 10003 (www address - http://www.thomson.com/chaphall/epd/epd-home.html). The CD-ROM single user is $250. The CD-ROM and book cost $280.

CI970343P

S0095-2338(97)00343-0