Variations in Medical Subject Headings (MeSH) mapping: from the natural language of patron terms to the controlled vocabulary of mapped lists*

By Lora V. Gault, M.Ed., M.S.
gault@calumet.purdue.edu
Instructional Services Librarian

The Library
Purdue University Calumet
169th Street and Wicker Avenue
P. O. Box 2590
Hammond, Indiana 46323-2590

Mary Shultz, M.S.
shultz@uic.edu
Assistant Health Sciences Librarian

Library of the Health Sciences-Urbana
University of Illinois at Chicago
102 Medical Sciences Building
506 South Mathews
Urbana, Illinois 61801

Kathy J. Davies, M.L.S.
kdavies@mail.mcg.edu
Education Services Coordinator

Greenblatt Library
Medical College of Georgia
Augusta, Georgia 30912-4400

Objectives: This study compared the mapping of natural language patron terms to the Medical Subject Headings (MeSH) across six MeSH interfaces for the MEDLINE database.

Methods: Test data were obtained from search requests submitted by patrons to the Library of the Health Sciences, University of Illinois at Chicago, over a nine-month period. Search request statements were parsed into separate terms or phrases. Using print sources from the National Library of Medicine, each parsed patron term was assigned corresponding MeSH terms. Each patron term was entered into each of the selected interfaces to determine how effectively they mapped to MeSH. Data were collected for mapping success, accessibility of MeSH term within mapped list, and total number of MeSH choices within each list.

Results: The selected MEDLINE interfaces do not map the same patron term in the same way, nor do they consistently lead to what is considered the appropriate MeSH term.
Conclusions: If searchers utilize the MEDLINE database to its fullest potential by mapping to MeSH, the results of the mapping will vary between interfaces. This variance may ultimately impact the search results. These differences should be considered when choosing a MEDLINE interface and when instructing end users.

INTRODUCTION

Users search online databases to find documents that “contribute to the satisfaction of some information need” [1]. The optimal result of a search is some compromise between the highest precision desired and the greatest recall possible to retrieve relevant documents. While searchers may use keyword approaches, a database-controlled vocabulary “has a very direct influence on system recall and precision capabilities” [2]. The controlled vocabulary for MEDLINE is the Medical Subject Headings (MeSH), produced by the National Library of Medicine.

Librarians often teach searchers how to use MeSH as a tool to find relevant articles indexed in the MEDLINE database. Using MeSH rather than simple keyword searching provides a greater level of access to article content beyond the terms used in titles and abstracts. MeSH indexing allows the specific content of articles to be clearly and consistently defined by controlled vocabulary terms. By using MeSH, indexers are able to define the same concept by multiple approaches and then assign headings to articles that increase both specificity and accessibility [3].

When searchers use MeSH, they are freed from having to account for synonyms and terminology variations that may be employed in indexed articles [4]. However, end users may find using the MeSH tools provided by online search systems to be a frustrating experience [5]. Searchers are often unaware of the MeSH terms assigned to specific concepts and rely on the search system to map entered terms to the correct headings. If exact MeSH terms are not found, online search interfaces often provide listings of subject headings from which users may choose an appropriate term. Currently, a variety of online vendors provide access to the MEDLINE database, and many offer the ability to search MeSH, however the process used to search MeSH is not identical among interfaces.

The differences that exist in each interface may impact the effectiveness of searching using MeSH mapping features. Several studies have provided evaluations of and comparisons among various interfaces to MEDLINE [6–8], but a comparison of the effectiveness of mapping natural language terms to MeSH across multiple MEDLINE interfaces was not found in the literature. Hallett’s study, comparing MeSH searches in DIALOG and OVID, revealed that the systems did not retrieve identical sets of documents. Hallett concluded that information professionals could not assume symmetrical retrieval from different online search systems when using the MeSH controlled vocabulary feature [9]. Lexical variants might also play a role in retrieving relevant search terms or results. Users might assume that any lexical variants such as word order or plural endings would be compensated for by online search systems. A comparison of the lexical variants of “bloody nose” revealed the fallacy of this assumption [10].

End users frequently enter non-MeSH terms for initial searches. The ability of each online system to correctly map users’ natural language terms to MeSH headings is crucial in achieving search precision and recall. A comparison of Internet Grateful Med (IGM), OVID, and PubMed reveals the unique features and functions of each interface with regard to searching and MeSH access [11]. De Groote notes that each system provides a method for suggesting MeSH terms, but the method of retrieving term choices varies greatly between the systems.

The present study compared the mapping of natural language terms to the controlled vocabulary of MeSH across six MEDLINE interfaces to determine if mapping was performed consistently among different MEDLINE search systems. The results obtained from a preliminary study conducted in 1999 indicated that the use of an identical patron term did not lead to the same MeSH term in each of the tested interfaces [12]. The investigators pursued this finding with an expanded study, drawing on search phrases and terms provided by patrons over a nine-month period.

METHODOLOGY

Test data were obtained from entries on search request forms submitted to the library at a comprehensive health sciences center that is part of a research university, between April 1998 and December 1998. These forms were used by patrons seeking to have mediated searches performed by librarians. The patrons were primarily composed of faculty, students, and staff from the colleges of nursing, medicine, pharmacy, public health, dentistry, and health and human development. Requests were also submitted by users from
other university departments as well as unaffiliated health professionals and consumers.

The forms were either filled out by requesters or by librarians during initial reference interviews with the requesters. The search requests were generally written in brief sentences or fragments. Examples include:

- influence of stress on toxicity of chemoprevention agents
- mitral valve prolapse—future risks and complications
- relinquishment in adoption—grief of birth parent

From the request forms, the statements were parsed into separate terms or phrases. Parsing of every request was done independently by each of the three investigators performing this study. The investigators then met and compared results, adding different terms and phrases to arrive at the broadest set of natural language terms. These terms encompassed a wide range of health care issues including conditions, treatments, drugs, administrative issues, social and behavioral issues, and managed care topics. The entire set of parsed patron terms was entered into a spreadsheet, and duplicate terms were removed. The final set of test data included 466 unique terms.

The search requests and their parsed patron terms were then divided between two of the investigators, who used the following sources from the National Library of Medicine as well as their own knowledge of medical terminology to assign the most appropriate corresponding MeSH term:

- Permutated Medical Subject Headings, 2000 [13]
- Medical Subject Headings—Annotated Alphabetic List, 2000 [14]
- Medical Subject Headings—Tree Structures, 2000 [15]

For the purposes of this study, the investigators chose to assign MeSH terms, excluding subheadings. Of the 466 original patron terms and phrases, 432 were found to have corresponding MeSH terms. Approximately 7% of the patron terms (34 terms) could not be associated with MeSH terms. In actual search situations, the investigators would have restricted these searches to the text or keyword fields and, accordingly, these terms were removed from the set of data used to test MeSH-mapping ability. The resulting test set was composed of 432 patron terms, each with its own assigned MeSH term. Examples of assigned terms are shown Table 1.

Each of the 432 patron terms and phrases were entered into each of the six selected interfaces to see how they mapped to MeSH. Each patron term was entered into each interface in the method described below, and the number of terms retrieved was recorded.

**Internet Grateful Med**

The Internet Grateful Med (IGM) search interface provided access to the Unified Medical Language System (UMLS) metathesaurus in addition to searching MeSH [16]. The UMLS retains term attributes present in MeSH and establishes synonymy and additional relationships between terms from MeSH and more than sixty additional biomedical vocabulary sources. The metathesaurus utilizes a single concept for each meaning and links to all the possible names in any of the vocabulary sources. For example “atrial fibrillation” is a concept including “auricular fibrillation” as well as the plural forms [17]. Patron terms were entered into the first search query box and the Find MeSH/Meta Terms button was selected to produce a list of MeSH terms. The investigators recorded the total number of terms displayed in the resulting list.

**PubMed Medical Subject Headings (MeSH) Browser**

The MeSH Browser can be selected from the sidebar menu in PubMed and provides users with the option to search MeSH for matching concepts to the entered terms or phrases. The PubMed MeSH browser provides a list of similar concept terms if an exact match is not found, but additional references from UMLS cannot be viewed in the MeSH Browser [18]. Each patron term was entered into the MeSH Browser and the GO button was selected. The result was a listing of terms inside a scroll box. The investigators scrolled to the end of this display box and recorded the total number of terms listed.

**PubMed Index/Preview feature**

The Index/Preview feature searches MeSH for an exact or partial alphabetical match to the entered term. If an exact match is not found, a list of alphabetical terms based on the first word is produced [19]. After choosing the Preview/Index option from the initial PubMed page, the MeSH Terms field was selected from the All Fields pull-down menu. Patron terms were then entered in the query box and the Index button was selected, which resulted in a displayed scroll box with a list of index terms. The Down button allows for continuous scrolling of the index, however, the investigators defined the bottom of the scroll box as the end-

<table>
<thead>
<tr>
<th>Parsed patron term</th>
<th>Assigned corresponding MeSH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prozac</td>
<td>Fluoxetine</td>
</tr>
<tr>
<td>QOL</td>
<td>Quality of life</td>
</tr>
<tr>
<td>Withstand pain</td>
<td>Pain threshold</td>
</tr>
<tr>
<td>Heard data</td>
<td>Auscultation</td>
</tr>
<tr>
<td>EDTA</td>
<td>Edetic acid</td>
</tr>
<tr>
<td>Paramedic</td>
<td>Emergency medical technicians</td>
</tr>
<tr>
<td>Anxiety</td>
<td>Anxiety</td>
</tr>
<tr>
<td>Prematurity</td>
<td>Infant, premature</td>
</tr>
</tbody>
</table>

**Table 1**

Examples of patron terms and their corresponding assigned Medical Subject Headings (MeSH) terms

Variations in MeSH mapping
point for counting terms. Thus, the total number of terms for this interface included only the terms listed in the initial scroll box.

OVID

The OVID search interface checks the “Tree or Thesaurus for the most appropriate terms and displays those terms in a list” [20]. The full MEDLINE file (1966–present) was used, and patron terms were entered into the initial query box. The Map Term to Subject Heading box was checked, and Perform Search was selected, which resulted in a list of MeSH terms. OVID lists the original search term as a final choice in each mapped list to address the possibility of keyword searching. Therefore, the investigators excluded the last item displayed in each mapped list when recording the total number of terms.

OCLC FirstSearch Index option using MeSH Heading Phrase

The OCLC Index search interface browses alphabetical lists of selected fields [21]. The investigators used the Index option, and MeSH Heading Phrase was chosen as the Indexed In field. Patron terms were entered into the query box, and the Browse button was selected. The system provided a list of terms and phrases in alphabetic proximity to the entered term, displayed in a table. There was an option to page forward for additional terms, which the investigators did until no new terms appeared. The total number of terms was determined and recorded by counting the terms listed on each individual page.

OCLC FirstSearch Index option using MeSH Heading

As noted above, the system browses MeSH for an alphabetical match to the entered term. In the Index option, MeSH Heading was selected as the Indexed In field. Patron terms were entered in the query box, and the Browse button was selected. The system produced a list of single-word terms displayed in a table. This list was based on the alphabetic position of the entered term and never included multi-word phrases, unless hyphenated. This interface also included an option to page forward for additional terms, which the investigators did until no new terms appeared. Similar to the OCLC FirstSearch MeSH Heading Phrase interface, the total number of terms was determined and recorded by counting terms listed on each individual page.

Measuring MeSH

The following questions were used to collect data for each term entered into each of the interfaces.

1. Is the corresponding MeSH term present?
2. If present, is it the first term?

3. If present, does it appear on the first screen?
4. If present, how far from the top of the list does it appear?
5. What is the total number of MeSH choices provided to users?

This method was used to determine if the corresponding MeSH term appeared in each interface and how close or accessible the term was to the user. The total number of choices available in each mapped list was also tabulated.

RESULTS

Each of the 432 patron terms was tested in each of the six MEDLINE interfaces to demonstrate the variance between the interfaces and to determine with which system users would have the most accurate MeSH mapping.

Mapping success

The first test question, “Is the MeSH term present?,” addresses the issue of success rates in mapping to MeSH. The results for each interface are presented in Table 2 and demonstrate discrepancies in the mapping performances of MEDLINE search interfaces. The scores for successfully mapping MeSH terms range from IGM’s high score of 72% to the lower performances (less than 50%) of the alphabetic mapping used by the PubMed Index option and both OCLC FirstSearch MeSH options. A discussion of mapping failures is provided in the “Failure analysis” section of this paper.

The specific success rates for MeSH mapping ob-
tained in this study varied from the pilot study [22],
but the rank order of the individual interfaces and
their relative positions within the ranking paralleled
that of the pilot study. In the pilot study, IGM scored
the highest with a success rate of 88%, followed by
PubMed with 72%, OVID at 65%, and then low scores
for the alphabetically mapped lists (less than 54%).
Given the results of this study and their correspond-
dence with the pilot study, it could be stated that the
interfaces did not map in the same way and that some
interfaces mapped to MeSH more successfully than
others. Although IGM had the highest score for map-
ping to the MeSH term in both the pilot study and in
the current study, this MEDLINE interface is no longer
available [23].

Accessibility

If the interface retrieved the assigned corresponding
MeSH term, then three additional questions were
asked: is the assigned term the first term in the list; is
the term on the first screen; and how far from the top
is the term?

These questions address the issue of accessibility. It
has been noted that “people tend to seek out inform-

Table 4
Corresponding MeSH on first screen of mapped list

<table>
<thead>
<tr>
<th>Search interface</th>
<th>Number on first screen per total mapped</th>
<th>Percent on first screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGM</td>
<td>249/309</td>
<td>81%</td>
</tr>
<tr>
<td>OVID</td>
<td>255/266</td>
<td>96%</td>
</tr>
<tr>
<td>PubMed MeSH Browser</td>
<td>258/271</td>
<td>95%</td>
</tr>
<tr>
<td>PubMed Index</td>
<td>200/208</td>
<td>96%</td>
</tr>
<tr>
<td>FirstSearch MeSH Heading Phrase</td>
<td>184/200</td>
<td>92%</td>
</tr>
<tr>
<td>FirstSearch MeSH Heading</td>
<td>141/141</td>
<td>100%</td>
</tr>
</tbody>
</table>

The issue of MeSH location in the mapped lists (Ta-
ble 5) was compared by noting the position of each

Table 5
Position in mapped list of corresponding MeSH term

<table>
<thead>
<tr>
<th>Search interface</th>
<th>Total number of mapped terms</th>
<th>1–5</th>
<th>6–10</th>
<th>11–15</th>
<th>16–20</th>
<th>21–25</th>
<th>26–30</th>
<th>31+</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGM</td>
<td>309</td>
<td>79%</td>
<td>6%</td>
<td>5%</td>
<td>1%</td>
<td>—</td>
<td>1%</td>
<td>8%</td>
</tr>
<tr>
<td>OVID</td>
<td>266</td>
<td>93%</td>
<td>4%</td>
<td>—</td>
<td>—</td>
<td>2%</td>
<td>—</td>
<td>1%</td>
</tr>
<tr>
<td>PubMed MeSH Browser</td>
<td>271</td>
<td>77%</td>
<td>19%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>—</td>
</tr>
<tr>
<td>PubMed Index</td>
<td>208</td>
<td>97%</td>
<td>—</td>
<td>—</td>
<td>1%</td>
<td>—</td>
<td>—</td>
<td>1%</td>
</tr>
<tr>
<td>FirstSearch MeSH Heading Phrase</td>
<td>200</td>
<td>92%</td>
<td>5%</td>
<td>2%</td>
<td>1%</td>
<td>—</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>FirstSearch MeSH Heading</td>
<td>141</td>
<td>94%</td>
<td>6%</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 3 documents the success of each interface in
displaying the corresponding MeSH term in the first
position in the list of choices provided to the user. The
PubMed Index option had a low success rate for map-
ping to the corresponding MeSH term, but of the terms
that were accurately mapped (208 of 432), most
were found to be at the top of the list (94%). PubMed
Index uses an alphabetic browse function so if the
browser finds a match, it is listed first and additional
choices in the list follow the patron term alphabetically.
Both FirstSearch options also use an alphabetic
mapping, but they scored the lowest for all interfaces
(16%). This performance can be attributed to the spe-
cific characteristics of FirstSearch’s alphabetical
mapping. The system compares patron terms to alphabeti-

Variations in MeSH mapping
term within the mapped list. Of the terms that mapped successfully, each interface performed well in displaying the corresponding MeSH term in the top ranges (within the first ten terms of a list). If the corresponding MeSH term was present, it was usually displayed (at least 75% of the time) within the first five entries of the mapped list for all of the selected interfaces. The presence of the corresponding MeSH term in the lower ranges of the mapped lists decreased sharply and steadily in all interfaces with the exception of IGM. IGM consistently had the longest mapped list and, similar to the other interfaces, the corresponding MeSH term was most frequently displayed within the first five entries of the list (79%). However, the investigators noted an increase in the appearance of the assigned MeSH term in the 31+ range for IGM that did not occur in any other system.

**Number of choices**

The average number of choices within each mapped list was also recorded for each interface. Table 6 shows a wide variance among the selected interfaces. IGM had the highest average of displayed MeSH terms with seventy-one, while OVID had the lowest at eight. The PubMed Index interface contained an option to page down through the entire alphabet. The investigators adjusted for the possibility of unlimited scrolling or paging by choosing the end point at the very first scroll box, which always contained forty terms.

The length of the list is easy to compare, but more complex issues arise when analyzing the results. While longer lists may, at first, appear to increase the likelihood of mapping to the appropriate subject heading, a closer look at the results reveals some inconsistencies. IGM had the highest mapping success rate (72%), but users would have to cope with an average of seventy-one retrieved terms. OVID, on the other hand, consistently provided the shortest list of choices to the user (average of eight terms), but their success rate was 62%, nearly identical to the PubMed MeSH Browser (63%), although the MeSH Browser provided lists nearly three times as long (average of 23 terms).

Is there an optimum length of choices for a list of subject headings? Lengthy lists may yield a more time-consuming search, requiring greater user patience and persistence. Wiberley et al. published a series of studies on user persistence in scanning lists of postings in online public access catalogs that suggested the ideal length may be thirty to thirty-five choices before adding further system options for user assistance [26–28]. The optimal length of choices for MeSH terms is an area requiring further study.

**Failure analysis**

There are several reasons why patron terms may not have mapped to the assigned MeSH terms. As noted earlier, the interfaces that used only alphabetic mapping were the least effective in mapping. The discussion in this section will focus on the three concept-based mapping interfaces: IGM, PubMed MeSH Browser, and OVID.

In some instances, the syndetic structure of the MeSH vocabulary itself seemed to be lacking. The syndetic structure of MeSH is represented by the tree structure numbers (not treated in this study) and the cross-references (such as the “consider also” cross-references, “forward see related” cross-references, “backward see” cross-references, etc.) [29]. This structure appears with the main headings and entry terms in the printed MeSH vocabulary. In other cases where the patron terms failed to map to the corresponding MeSH terms, the structure was present in the printed vocabulary but was not apparently utilized by the online interface. In a few cases, the patron terms were so unusual and could not possibly be accounted for in either the vocabulary or in the mapping algorithm used by the interfaces.

Some of the failures could be attributed to weaknesses within the MeSH vocabulary. For example, the patron term “ingestion” was assigned the MeSH term “eating.” Ingestion did not successfully map in any of the interfaces. Although ingestion would be a likely candidate as an entry term (“see” cross-reference term) within MeSH, it did not appear in the MeSH print sources. The patron term “cell phones” was not located in the print sources but was assigned the MeSH term “telephone” by the investigators. Cell phones would seem likely as a candidate for a MeSH term itself or as a cross-reference from telephones. Only the OVID interface mapped this term successfully, while both IGM and the PubMed MeSH Browser mapped cell phones to terms related to cells or cellular structures.

Other mapping failures would not have occurred if the interface fully utilized the syntetic structure already available in MeSH. Some of the patron terms that did not map to corresponding MeSH terms were, in fact, located in the “Medical Subject Headings—Annotated Alphabetic List” as entry terms (see cross-reference terms) or as “see” references. For example,
when looking for the patron term “paramedic,” the following see references were found:

**Paramedical Personnel** see Allied Health Personnel
**Paramedics** see Allied Health Personnel
**Paramedics, Emergency** see Emergency Medical Technicians

The MeSH term “emergency medical technicians” was assigned to this patron term. However, when paramedic was entered into both the OVID interface and the PubMed MeSH Browser interface, “allied health personnel” appeared in the mapped lists of terms, but emergency medical technicians did not. This indicates that perhaps these two interfaces were not utilizing the full functions and structure of the MeSH vocabulary.

The apparent use of an online permuted system for mapping worked in some instances. The patron phrase “withstand pain” was assigned the MeSH term “pain threshold,” using the Permutated Medical Subject Headings. Withstand pain was successfully mapped to pain threshold in both IGM and the PubMed MeSH Browser but failed in the OVID Interface.

In some cases where patron terms did not have a direct match to MeSH entry terms (see cross-reference) or “see related” reference within the print sources, OVID’s automatic mapping succeeded where both IGM and the PubMed MeSH Browser failed. For example, the patron term “GI tract” mapped to the assigned MeSH term “gastrointestinal system” in OVID but not in IGM or the PubMed MeSH Browser. The same was true of the patron term “QOL” and its corresponding assigned MeSH term “quality of life.”

A few of the mapping failures were understandable in terms of the patron phrases. For example, the patron phrase “heard data” was assigned the MeSH term “auscultation,” and the patron phrase “large dosages” was assigned the MeSH term “maximum tolerated dose.” These assignments were made based on the context of the entire search request statement and the investigator’s own knowledge and experience with MeSH. These patron phrases did not map to the assigned MeSH terms in any of the interfaces. The MeSH vocabulary could not reasonably include all possible natural language terms as entry terms (see cross-reference terms). There will always be cases where the skills of search intermediaries such as medical librarians would be useful.

IGM had the highest mapping success rate (72%) and utilized the UMLS metathesaurus. OVID and the PubMed MeSH Browser were close in their success rates (62% and 63%), however, OVID used its own algorithm for mapping to MeSH, which was often successful in cases where both IGM and the PubMed MeSH Browser failed. Combining the finest qualities of the mapping within each interface would be useful.

**CONCLUSION**

The selected MEDLINE search interfaces do not map the same patron terms in the same way, nor do they consistently lead to what is considered the appropriate MeSH term. If searchers use the MEDLINE database to its fullest potential by mapping to MeSH terms, the results of this mapping will vary among different interfaces and not always retrieve the appropriate MeSH term. This study suggests that the choice of a MEDLINE interface and the resulting mapping options are critical factors in the success of searches. The differences in MeSH mapping should be considered when choosing a MEDLINE search interface and when instructing end users.

The investigators noted that IGM showed the highest number of terms mapped (72%), but this interface to MEDLINE has been retired. IGM used the UMLS metathesaurus to make connections between patron terms and subject heading concepts. The inclusion of additional vocabulary sources in UMLS might have played a role in the higher success rate for mapping terms to MeSH. The results from both the pilot and the expanded study indicated that it might be advisable to incorporate some of the mapping mechanisms from IGM into other MEDLINE interfaces, such as the PubMed MeSH Browser.

The results of this study also demonstrate that mapping natural language terms to subject headings based solely on the alphabetic position of the entered term is ineffective. For example, “heart attack” will map to “heart auscultation,” but not “myocardial infarction.” For the greatest accuracy, mapping from natural language to controlled vocabularies should be concept-based, not alphabetical.

**ACKNOWLEDGMENTS**

The authors would like to give special thanks to the following individuals for their assistance and encouragement throughout the course of this project: Ann Weller, Stephen Wiberley, and Victoria Pifalo of the University of Illinois at Chicago and Pauline Cochrane of the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign.

**REFERENCES**

2. Ibid., 108.
19. Ibid.
22. Davies, op. cit.
23. Knecht L, Shooshan S. Internet Grateful Med to be retired; reminder of NLM Gateway availability. NLM Tech Bull 2001 Jan-Feb;318(e3).

Received October 2001; accepted November 2001