Improving PubMed Search Using Medical Subject Heading (MeSH) Terms

Mahvish Umar [1]*, Sadaf Sharif [2], Sana Younas [1]

Abstract— Medical Subject Headings (MeSH) terms are valuable tools to conduct focused and concise search. This review article evaluates the performance of search strategies for retrieval of information from the national Library of Medicine (NLM) by using the Medical Subject Headings (MeSH) while searching for articles about analgesics used for pain management in cancer patient. We limited our search to 10 types of analgesics. We compared two search strategies MeSH terms and text word searching. Results emphasized that search through MeSH terms provide more precise, efficient and time saving information compared to the text word search. Understanding and learning how to search in PubMed using MeSH terms strategy tool is highly recommended for healthcare professionals to efficiently retrieve the search related information in a timely manner.

Keywords: Key words: PubMed (MeSH), Medical Subject heading (MeSH), Pain (MeSH), Cancer Pain (MeSH), Drug Therapy (MeSH), Analgesics (MeSH).

1. INTRODUCTION

To better manage and search the biomedical literature, the US National Library of Medicine (NLM®) developed the Medical Subject Heading (MeSH) controlled vocabulary for indexing MEDLINE® articles. MeSH has been used to improve PubMed query results [1,2], information retrieval, document clustering, and query refinement in 'downstream' applications that use PubMed® abstracts. For instance, Abdou and Savoy [3] have shown that including MeSH can improve the retrieval performance from 2.4% to 13.5% among 10 different information retrieval models. MeSH has also been used in several web tools for searching

biomedical literature to cluster documents [4]. Furthermore, visualization tools like LigerCat [5] and MeSH [1] have used MeSH terms to help users understand and filter PubMed query results.

Alternatively, PubMed employs a process called Automatic Term Mapping (ATM) that compares and maps untagged terms from the user query to lists of pre-indexed terms in PubMed's translation tables in the following order: The Medical Subject Headings (MeSH) table (mapping search terms to MeSH concepts), the journals translation table (mapping search terms to journal names), and the author index (mapping search terms to author names). In the context of query expansion, we considered only the translation via MeSH, which is designed by the NLM for indexing and searching of the MEDLINE database of journal citations. So, if a user query includes a term that can be mapped to a MeSH concept, the search will then add the MeSH term to the original query.

As a result, the ATM process enables the original query to have an access to the MeSH field of the indexed MEDLINE documents. For instance, if a user submits a query tumor with no search tags attached, the query will be automatically mapped to the MeSH Term Neoplasms by ATM. Thus, not only documents having the word tumor in the title and abstract will be retrieved, any other documents indexed with the MeSH term Neoplasms (as well as more specific terms beneath it in the MeSH hierarchy) will be returned as well. Furthermore, individual words in MeSH concepts are also indexed and searched. Taking the previous query as example, documents indexed with MeSH terms that include the word tumor (e.g. Tumor Virus Infections) will also be returned.

By default, PubMed applies the Automatic Term Mapping feature to every query. It is designed and implemented by the NLM as a means to improve retrieval performance. Unlike the studies in the past that focused on the accuracy of the mapping process [6-9]. The aim of this review is to investigate the performance of search strategies in the retrieval of indexed information by using MeSH terms compared to key words for different types of analgesics used for pain management in cancer patient.

[1] Gujranwala Institute of Nuclear Medicine and Radiotherapy, Gujranwala, Pakistan.
[2] Rashid Latif Medical Complex, Gujranwala, Pakistan

*Corresponding Author: Mahvish Umar, PharmD
Email address: mahvishumar96@gmail.com

2. Materials & methods

We compared searches for ten randomly selected different analgesics used to manage cancer pain by comparing the number of retrieved articles for both MeSH terms and text word searches, with the assumption that the search obtaining the fewer number of articles was considered more efficient way to initiate a search. The opioid analgesics used were codeine, dihydrocodeine, morphine, tramadol, methadone, hydrocodone, fentanyl, oxymorphone, oxycodone and dihydrocodeine. The number of articles found by each search strategy was noted, and the percentage difference was calculated by dividing the difference in article retrievals by the number of retrieved text word articles. We also conducted an in-depth review of one search and studied the abstracts of articles found by the text word search strategy but missed by MeSH. Articles that showed relevance based on their abstract alone went through further evaluation of the full text. We then concluded whether these articles were pertinent to the aim of the search. The number of relevant articles missed by MeSH was recorded, and the percentage of missed articles was calculated by dividing the number of missed articles by the number of Medical Subject missed relevant articles. Search queries in PubMed can be made with MeSH. In text word search PubMed scans the whole record of articles, titles, abstracts, list of applied MeSH terms, list of authors and journal name. Articles containing queried text words in any one of these areas are retrieved. As a result, text word searches often identify articles not relevant to the search topic. For example, a text word search on the "Analgesics used for pain management in cancer patients" would not only identify the related articles, but would list every article with the word Analgesic or pain management. Additionally, all possible synonyms and variants of terms must be specified in order to ensure an inclusive result. Search terms were entered into the query box on the main (home) PubMed search page and then searched. The search details page was checked to determine how PubMed processed the search [10].

Specifically, we were interested in determining how or whether the search term was translated or mapped to a MeSH term. For example, for analgesics we used the clinical question to provide a context for judging whether a search term mapped to an appropriate MeSH term, the clinical question is "what analgesics used for cancer pain?" by adding this question we didn't find any article in PubMed then we selected NCBI literature resource MeSH down from screen, and added MeSH term "Pain" in the search box then we selected the word "cancer pain" from hierarchy and "drug therapy" from subheadings under this MeSH term. we added these terms to search builder and search articles from PubMed. After studying these articles, we found that opioid analgesics are used for cancer pain. Then after we went back to the PubMed search engine and added the term. "Analgesics" to the search box and found "Opioid analgesics" in hierarchy and search articles. Then we found the articles which are quite relevant to our research requirements. While with the text word search, we found that articles of cancer as well as articles of pain and articles for analgesics were included. If a search term automatically mapped to a MeSH term, the definition of the MeSH term was reviewed to determine whether the MeSH term conceptually related to the clinical question. We also examined the MeSH hierarchy to determine whether appropriate broader or narrower MeSH terms existed. Both the definition of the MeSH term and the MeSH hierarchy were accessed through the MeSH database. When a search term did not automatically map to a MeSH term, we entered the search term into the query box of the MeSH database to determine whether a MeSH term conceptually related to the clinical question was available. If so, we again examined the MeSH hierarchy to determine whether appropriate broader or narrower MeSH terms existed.

3. results

We reviewed ten opioid analgesic topics and compared the results using MeSH and text word terms. We searched for the analgesic's names and topics. The number of articles retrieved are displayed in Table I and Figure I. The opioid analgesics used are codeine, dihydrocodeine, morphine, tramadol, methadone, hydrocodone, fentanyl, oxymorphone, oxycodone and dihydrocodeine. MesH search was more effective in redeeming the concerned articles.

Figure 2 shows sampling of Mesh headings pertinent to cancer pain including the MESH terms for analgesics. The Mesh headings and categories are listed to demonstrate the hierarchy of terms used to show how MeSH terms are arranged. Figure 3 shows a very limited sampling of the MeSH tree for analgesics.

MeSH searches retrieved fewer journal articles. Simple, but not otherwise quantitated, review of each search showed that the text word searches had more extraneous articles than the better focused MeSH searches. The MeSH strategy retrieved 245 articles, compared with 982 articles retrieved by the text word strategy. We then reviewed the 737 articles that were not found by the MeSH search. Of the articles not retrieved by MeSH, only 12 articles (both from the 1970's to 2023) were selected for further examination based on a review of their abstracts. After reading the full text, articles were determined to be relevant to the search. Therefore, 1.6% of the articles missed were articles pertinent to the search. Of note, the majority of the irrelevant articles retrieved by the text word search included articles which focused on various conditions, such as administration and dosage only. When the "Detail" tab was selected to show the search algorithm of how the articles were retrieved, many of these articles were found because the word "administration" was associated in the title or abstract. Searches can be further modified by using limits which are easily applied using the "Limits" tab under the search box on the main PubMed page. These will limit searches by language, age group, gender, publication type, etc [11,12].

4. DISCUSSION

Whenever we need information related to indexed articles for biomedical search we consult PubMed search engine, but

Table 1: Comparison between the percentage of search topic using MeSH terms versus text search

Search Topics		Number of Articles	% difference
Administration and dosage of tramadol Tramadol administration/dosage	(Text word) (MeSH)	1851 1221	34%
Adverse effects of Hydromorphone Hydromorphone /adverse effects	(Text word) (MeSH)	814 328	59%
Analogs and derivatives of Methadone Methadone Analog/Derivatives	(Text word) (MeSH)	811 339	58%
Supply and distribution of Morphine Morphine /supply and distribution	(Text word) (MeSH)	333 42	87%
Toxicity of oxycodone Oxycodone /Toxicity	(Text Word) (MeSH)	220 23	89%
Administration and dosage of Hydrocodone Hydrocodone/Administration and dosage	(Text word) (MeSH)	362 203	44%
Antagonist and Inhibitors of Oxymorphone Oxymorphone /Antagonists and inhibitors	(Text Word) (MeSH)	74 14	81%
Therapeutic use of fentanyl Fentanyl /Therapeutic uses	(Text Word) (MeSH)	20162 9999	50.4%
Metabolism of codeine Codeine / Metabolism	(Text Word) (MeSH)	3074 1421	54%
Adverse effects of Dihydrocodeine Dihydrocodeine /Adverse effects	(Text words) (MeSH)	3351 1841	45%
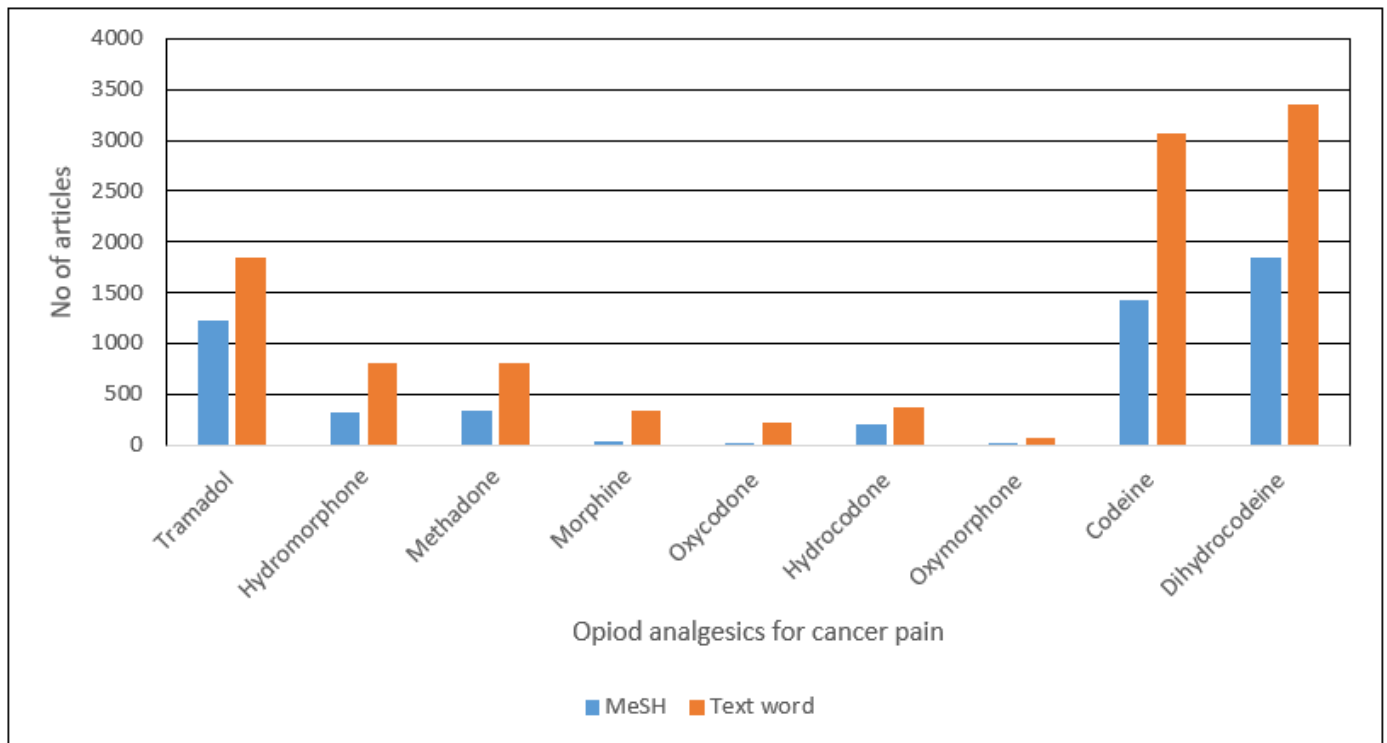
Figure 1: Differences in search results between MeSH terms versus text words
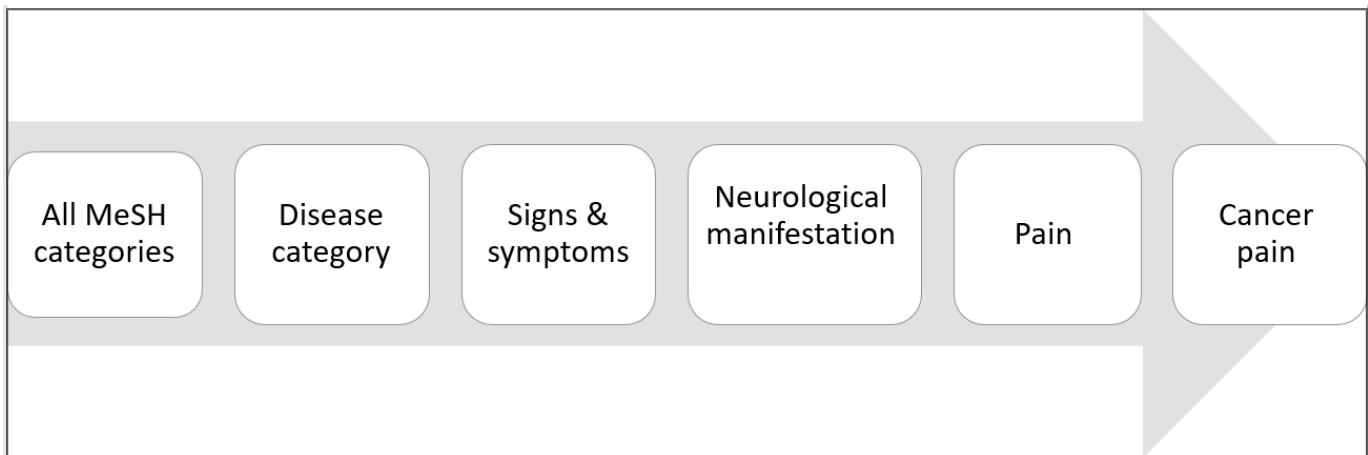
Figure 2: Samples for MeSH heading
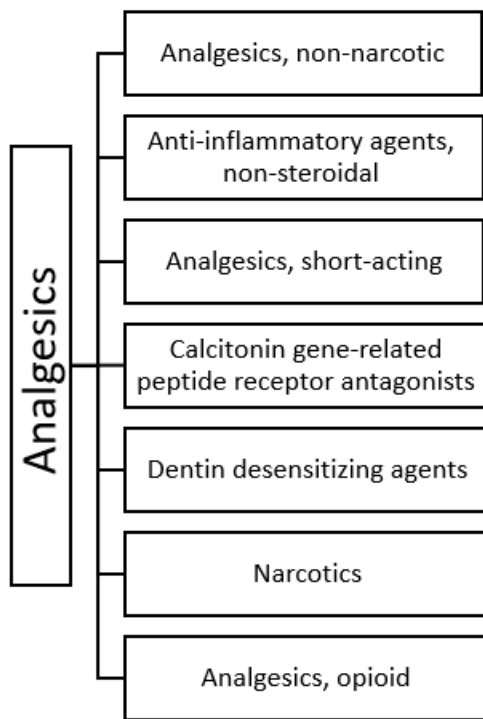
Figure 3: MeSH tree for analgesics

in order to save our time, we can use MeSH terms to focus on relevant and precise articles. A MeSH thesaurus is the best tool for improving access to rapidly growing biomedical knowledge. MeSH is organized as a complex of hierarchy called the "MeSH tree structure", in which the MeSH terms are arranged in a tree like branching structure that increases specificity. The original system was attached with multiple occasions but never retrieved. Hence as time passes additional categories and subcategories were added. The MeSH terms make PubMed content more easily accessible to health professionals and links MEDLINE users to publishers [13]. An example, cancer related entries per year in PubMed has risen from around 6% in 1950 to more than 16% 2016 [14].

In this paper we compared the performance of two search strategies in searching different aspects of 10 analgesics used to manage cancer pain by using text words search and MeSH terms search. We found that MeSH terms retrieve less, focused and relevant journal articles than the text word searches. In addition, we discovered that MeSH terms improve efficiency as many terms "mapped "to the MeSH term when one uses the PubMed MeSH database. After typing the search word, the system automatically links the preferred term. For example, in search analgesics for cancer pain we found articles for opioid analgesics such as morphine, codeine used to treat cancer pain while in text words we find articles which contain irrelevant information about all the analgesics available.

The comparison of MeSH search with text word search is not new. These comparisons are performed to help authors to utilize easy and focused search tools. Authors have compared literature searches by using MeSH terms and text words and evaluate the accuracy of outcomes on bases of their sensitivity and specificity. Sensitivity is defined as the ability of the search to retrieve the relevant citations from the database. It can also be viewed as the recall or the number of relevant citations retrieved divided by the number of relevant citations in the database. Specificity is the ability of searches to differentiate between the relevant and non-relevant citations. This can be viewed as the precision of the search. Precision is the number of relevant citations retrieved divided by the total number of citations retrieved. Therefore, high sensitivity of search signifies that important articles were identified. And high specificity indicates that review was efficient and few important articles were retrieved.

Studies show that text word searches have lower specificity than searches performed using MeSH. It was reported in a study comparing MeSH and text word searches to retrieve studies on sleep, that the MeSH search produced higher specificity than the text word search (66% and 47%, respectively), but lower sensitivity (78% for the MeSH search versus 88% for the text word search) [15]. Another, unpublished study of 975 MEDLINE searches at Harvard Medical School showed that MeSH searches provided significantly higher specificity and sensitivity than title-abstract text word searches [7]. Other authors have argued that

the sensitivity and specificity of searches are enhanced when searches combine both text words and MeSH words while retrieving articles. In addition, the authors found that although the specificities of individual MeSH terms and text words were high, this was often at the expense of their sensitivity. When combining terms, however, the high specificity was maintained with modest increases in sensitivity [16].

Skills using PubMed to search the NLM is important to retrieve accurate, up-to-date medical information. It is highly recommended that researchers use MeSH tools to conduct their search. A drawback of MeSH and text word searches is that articles exhibit in a sequential manner. Therefore, it is suggested that once the search is regulated, one identifies the relevant articles and selects, "Related Articles". This triggers a second search using the MeSH terms assigned to the identified article and displays results in order of their relevance to the original article, as measured by similar MeSH and text word terms. In some occasions when a broadcast search is required and when a search involve unusual conditions then we have to use text word search**.**

4. CONCLUSION

Searching using PubMed MeSH terms is more efficient, concise and time saving compared to key words search. This search tool is highly recommended to be utilized while conducting a search to obtain the most relevant articles in a timely manner.

Conflict of Interest
The research was conducted without any commercial or financial relationships that can be of any potential conflict of interest.

Authors Contribution
The Authors confirms sole responsibility for the study conception and design data collection, analysis and interpretation of results and manuscript preparation.

REFERENCES

1. Lu Z, Kim W, Wilbur WJ. Evaluation of Query Expansion Using MeSH in PubMed. Inf Retr Boston. 2009;12(1):69-80. doi: 10.1007/s10791-008-9074-8.

2. Richter RR, Austin TM. Using MeSH (medical subject headings) to enhance PubMed search strategies for evidence-based practice in physical therapy. Phys Ther. 2012 Jan;92(1):124-32. doi: 10.2522/ptj.20100178.

3. Samir Abdou, Jacques Savoy, Searching in Medline: Query expansion and manual indexing evaluation, Information Processing & Management, Volume 44, Issue 2, 2008, Pages 781-789, https://doi.org/10.1016/j.ipm.2007.03.013.

4. Lu Z. PubMed and beyond: a survey of web tools for searching biomedical literature. Database (Oxford). 2011 Jan 18;2011:baq036. doi: 10.1093/database/baq036. PMID: 21245076

5. Sarkar, I. N., Schenk, R., Miller, H., Norton, C. N. LigerCat: using "MeSH Clouds" from journal, article, or gene citations to facilitate the identification of relevant biomedical literature. AMIA Annu Symp Proc. 2009:2009: 563-567.

6. Gault LV, Shultz M, Davies KJ. Variations in Medical Subject Headings (MeSH) mapping: from the natural language of patron terms to the controlled vocabulary of mapped lists. J Med Libr Assoc. 2002 Apr;90(2):173-80. PMID: 11999175; PMCID: PMC100762.

7. Carlin BG. PubMed automatic term mapping. J Med Libr Assoc. 2004 Apr;92(2):168. PMID: 15088076; PMCID: PMC385296.

8. Smith AM. An examination of PubMed's ability to disambiguate subject queries and journal title queries. J Med Libr Assoc. 2004 Jan;92(1):97-100. PMID: 14762470.

9. Shultz M. Mapping of medical acronyms and initialisms to Medical Subject Headings (MeSH) across selected systems. J Med Libr Assoc. 2006 Oct;94(4):410-4.

10. Yu Z, Bernstam E, Cohen T, Wallace BC, Johnson TR. Improving the utility of MeSH® terms using the TopicalMeSH representation. J Biomed Inform. 2016 Jun;61:77-86. doi: 10.1016/j.jbi.2016.03.013. Epub 2016 Mar 19.

11. Lowe HJ, Barnett GO. Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches. JAMA. 1994 Apr 13;271(14):1103-8. PMID: 8151853.

12. Coletti MH, Bleich HL. Medical subject headings used to search the biomedical literature. J Am Med Inform Assoc. 2001 Jul-Aug;8(4):317-23. doi: 10.1136/jamia.2001.0080317. Erratum in: J Am Med Inform Assoc 2001 Nov-Dec;8(6):597.

13. Lindberg DA. Internet access to the National Library of Medicine. Eff Clin Pract. 2000 Sep-Oct;3(5):256-60. PMID: 11185333.

14. Reyes-Aldasoro CC. The proportion of cancer-related entries in PubMed has increased considerably; is cancer truly "The Emperor of All Maladies"? PLoS One. 2017 Mar 10;12(3):e0173671. doi: 10.1371/journal.pone.0173671.

15. Jenuwine ES, Floyd JA. Comparison of Medical Subject Heading and text-word searches in MEDLINE to retrieve studies on sleep in healthy individuals. J Med Libr Assoc 2004;92(3):349-353.

16. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc. 1994 Nov-Dec;1(6):447-58. doi: 10.1136/jamia.1994.95153434.