Research in Social and Administrative Pharmacy 20 (2024) 911–917

Contents lists available at ScienceDirect

Research in Social and Administrative Pharmacy

journal homepage: www.elsevier.com/locate/rsap

Original Research

Influence of automated indexing in Medical Subject Headings (MeSH) selection for pharmacy practice journals

Fernando Fernandez-Llimos [a,*], Luciana G. Negrão [a], Christine Bond [b], Derek Stewart [c]

[a] Applied Molecular Biosciences Unit (UCIBIO), Laboratory of Pharmacology, Department of Drug Sciences, Faculty of Pharmacy, University of Porto, Porto, Portugal
[b] Primary Care, University of Aberdeen Institute of Applied Health Sciences, Aberdeen, United Kingdom
[c] College of Pharmacy, QU Health, Qatar University, Doha, Qatar

ARTICLE INFO

Keywords:
Medical subject headings
Terminology as topic
Publications
Periodicals as topic
Pharmacists
Pharmacy

ABSTRACT

Background: The Medical Subject Headings (MeSH) thesaurus is the controlled vocabulary used to index articles in MEDLINE. MeSH were mainly manually selected until June 2022 when an automated algorithm, the Medical Text Indexer (MTI) automated was fully implemented. A selection of automated indexed articles is then reviewed (curated) by human indexers to ensure the quality of the process.
Objective: To describe the association of MEDLINE indexing methods (i.e., manual, automated, and automated + curated) on the MeSH assignment in pharmacy practice journals compared with medical journals.
Methods: Original research articles published between 2016 and 2023 in two groups of journals (i.e., the Big-five general medicine and three pharmacy practice journals) were selected from PubMed using journal-specific search strategies. Metadata of the articles, including MeSH terms and indexing method, was extracted. A list of pharmacy-specific MeSH terms had been compiled from previously published studies, and their presence in pharmacy practice journal records was investigated. Using bivariate and multivariate analyses, as well as effect size measures, the number of MeSH per article was compared between journal groups, geographic origin of the journal, and indexing method.
Results: A total of 8479 original research articles was retrieved: 6254 from the medical journals and 2225 from pharmacy practice journals. The number of articles indexed by the various methods was disproportionate; 77.8 % of medical and 50.5 % of pharmacy manually indexed. Among those indexed using the automated system, 51.1 % medical and 10.9 % pharmacy practice articles were then curated to ensure the indexing quality. Number of MeSH per article varied among the three indexing methods for medical and pharmacy journals, with 15.5 vs. 13.0 in manually indexed, 9.4 vs. 7.4 in automated indexed, and 12.1 vs. 7.8 in automated and then curated, respectively. Multivariate analysis showed significant effect of indexing method and journal group in the number of MeSH attributed, but not the geographical origin of the journal.
Conclusions: Articles indexed using automated MTI have less MeSH than manually indexed articles. Articles published in pharmacy practice journals were indexed with fewer number of MeSH compared with general medical journal articles regardless of the indexing method used.

1. Introduction

There is an increasing body of work published in the discipline of clinical and social pharmacy practice (hereinafter referred to as pharmacy practice) much of which is relevant to the wider health care community. However, its accessibility may be limited if it is not optimally indexed using the Medical Subject Headings (MeSH) thesaurus. MeSH is the U.S. National Library of Medicine's (NLM) controlled vocabulary thesaurus, created in 1960 to standardize indexing of health-related literature.[1] MeSH terms facilitate the retrieval of scientific studies which address similar concepts although the study authors may have used different terminology in their titles or abstracts. MeSH terms are allocated to articles indexed in MEDLINE; PubMed, a search engine, retrieves citations from MEDLINE and other bibliographic databases

* Corresponding author. Laboratory of Pharmacology, Department of Drug Sciences, College of Pharmacy, University of Porto, R. Jorge de Viterbo Ferreira 228, 4050-313, Porto, Portugal.
E-mail addresses: fllimos@ff.up.pt (F. Fernandez-Llimos), up202310533@edu.ff.up.pt (L.G. Negrão), c.m.bond@abdn.ac.uk (C. Bond), d.stewart@qu.edu.qa (D. Stewart).

https://doi.org/10.1016/j.sapharm.2024.06.003
Received 10 June 2024; Accepted 10 June 2024
Available online 12 June 2024

based on these MeSH terms and authors' free text words.[2]

Using MeSH terms in search queries improves the efficiency and precision of the search, reducing the number of irrelevant citations retrieved, especially when inconsistent terminology is used by authors publishing in the same field. Ignoring MeSH terms in search strategies increases the complexity of the search (i.e., number of text words) needed to retrieve all the relevant studies.[3–5] Nonetheless, some issues have been identified with MeSH vocabulary. For example, there are fewer terms and less specific terms, when compared to the Scopus controlled vocabulary, the Emtree.[6] Also, deficient coverage of MeSH in some fields has been reported.[7,8] Pharmacy practice is an example of a poorly covered field. In 2014 the existence of only 26 pharmacy-specific terms is reported, while nursing and dentistry had 94 and 145, respectively.[7,9] Minguet et al. proposed 16 new pharmacy-specific MeSH terms, with two immediately included by the NLM,[10] and three more in the following years. The process of indexing articles with the existing MeSH terms was deemed inaccurate, especially in areas that may not be considered as core for medicine: acronyms used in the abstract not mapping to the appropriate term, or not allocation of MeSH terms that exactly match with words in the abstract.[8,11,12] Furthermore, the delay in allocating MeSH to articles (indexing) in pharmacy practice has been shown to be longer than in other medical areas.[13]

In 1996, the NLM started the Indexing Initiative project, aiming to explore new indexing technologies.[14] Until 2011, MeSH indexing was conducted entirely manually by NLM human catalogers. In 2011, recognizing the increasing volume of publications and the need to reduce indexing times, NLM introduced the algorithm Medical Text Indexer (MTI) to propose MeSH terms for 14 journals. These potential MeSH terms were then checked for relevance by humans. The number of journals indexed with support of the MTI increased to 230 in 2014. In 2019, fully automated indexing was initiated with a (unspecified) group of journals, using the MTI-Auto (MTIA), with human curation (i.e., checked for quality assurance) of selected articles. From April 2022, all journals in MEDLINE are entirely indexed by the MTIA. To ensure the quality of the process, NLM stated that human catalogers would continue to curate citations involving genes and proteins, cases of known ambiguity (terms with more than one potential meaning), clinical trials, and an additional random set of citations. Although the use of articles' full text was proposed for the purposes of indexing,[14] the MTIA uses only article's title and abstracts and the MeSH terms of PubMed related records to generate a list of MeSH terms, Supplementary Concept, and Publication Type descriptors selected for the article being indexed.[15] The effects of this shift from humans to automated processes on the quality of MeSH assignment has not been fully evaluated.[16] The objective of this study was to describe the association of MEDLINE indexing methods (i.e., manual, automated, and automated + curated) on the MeSH assignment in pharmacy practice journals compared with medical journals. Given that NLM is U.S.-based, the association of geographical region and MeSH assignment was also explored.

2. Methods

This was a longitudinal study comparing two groups of journals (i.e., pharmacy practice journals and general medical journals) during an 8-year period (four years pre- and four years after the full implementation of the MTIA).

2.1. Data collection

Journals were selected to represent two different groups: pharmacy practice and general medicine journals. Pharmacy practice journals were selected from those indexed in MEDLINE and which had adopted the Granada Statements (https://granadastatements.weebly.com/), namely the International Journal of Clinical Pharmacy, International Journal of Pharmacy Practice, and Research in Social and Administrative Pharmacy. General medical journals were those commonly known as the 'Big-five' medical journals,[17–19] which usually serve as a gold-standard of best medical publishing practices, namely, Annals of Internal Medicine, British Medical Journal, Journal of the American Medical Association, The Lancet, and the New England Journal of Medicine.

Since no standard terminology exists to describe an article as an original research article, and since these articles are not always cataloged as such in bibliographic databases, a systematic search was designed to retrieve original research articles published in each of the selected journals. An article was considered as an original research article if it included a structured abstract with discrete sections for methods and results. As terminology used in the abstract structure is not consistent across journals, search strategies were adapted to the terms used in each journal. Terms such as Methods, Measurements, Design, Setting, or Participants were used for the methods section. Terms such as Results, or Findings were used for the results section. The specific search strategies for each journal are described in Supplementary Material.

Specific searches for each journal were executed in PubMed (April 15, 2024) using these search strategies. Retrieved records were exported into TXT files using two different procedures:

- A PUBMED format file was obtained using the "Save citations to file" feature of PubMed. This file contained the complete record, including the following fields of interest: the MeSH terms, the MDAT (date of MeSH assignment), the EDAT (entry date in PubMed), and the DP (date of publication).
- A PMID format file was created using the "Save citations to file" feature of PubMed. This file comprised the list of PMIDs of the articles retrieved for each journal.

To obtain the XML record, which included the indexing method recorded for each article, the PMID list was used in the PubMed2XML tool (https://pubmed2xl.com/xml/). The attribute ≪IndexingMethod≫ in the ≪MedlineCitation≫ element was extracted for each article, when available.

For the purposes of this study, a list of pharmacy-specific MeSH terms was created based on the original list suggested by Minguet et al.[12] with the addition of the MeSH created as a consequence of that initial study.[7,10] This generated a list of 31 pharmacy-specific MeSH.

2.2. Data analysis

A Microsoft Excel (Microsoft, Redmond, WA, United States) database with all the original research articles retrieved for each journal was created comprising all metadata obtained from PubMed exported files. All the MeSH terms attributed to each article were extracted for analysis.

Null hypothesis tests were used to determine association between categorical (i.e., chi square) or continuous (i.e., Student's *t*-test, ANOVA) variables. Significance was established at $p < 0.005$. Following American Statistical Association's recommendations,[20] effect size measures were also calculated (i.e., Cohen's d, eta-square, odds ratio - OR). A multivariate analysis through a linear regression was conducted with the number of MeSH as dependent variable and three covariates (i.e., geographic origin, group of journals, and indexing method). SPSS v25 (IBM, Armonk, NY, United States) was used for statistical calculations. Effect sizes were categorized into null, moderate, and strong, following Cohen's classification.[21]

3. Results

The eight journals under analysis published a total of 8479 original research articles during the eight years; 6254 published by the five big-five medical journals and 2225 by the three pharmacy practice journals. The number of articles retrieved was relatively consistent across years, with a mean of 1060 articles per annum (SD = 52; range 967-1114); 781 (SD = 22; range 741–806) published in the big-five medical journals and

278 per annum (SD = 60; range 171–335) in the pharmacy practice journals.

Overall, 5990 (70.6 %) articles were manually indexed, compared to 1659 (19.6 %) indexed using the automated algorithm alone and 830 (9.8 %) using the automated algorithm followed by curation (automated + curation). Medical journals had 29 articles (0.4 %) indexed using the automated algorithm between 2016 and 2021, increasing to 70.9 % in 2022, and full implementation in 2023. The first automated indexing among pharmacy practice journals was recorded in 2020 (49.7 %), increasing to 92.6 % in 2021, and 100 % in 2022 (Table 1).

The proportion of articles indexed with the different methods varied between the geographical origin of the journals (Table 2). While the proportion which were manually indexed was similar in U.S. journals compared with non-U.S. journals (70.1 % and 71.4 %, respectively), there was a statistically significant difference in the percentage of automated indexed articles that were then curated (chi-square $p < 0.001$; OR 7.21; 95%CI 5.79:8.97), with 48.3 % U.S. with additional curation out of those indexed by the automated algorithm but only 11.5 % non-U.S.

Considering journal groups, significant difference existed in the number of articles indexed with the automated algorithm between medical and pharmacy practice journals ($p < 0.001$; OR 3.43; 95%IC 3.10:3.80), with 49.5 % articles in pharmacy practice articles indexed using the automated algorithm and only 22.2 % articles in medical journal. Additionally, a significantly higher proportion of medical journals articles indexed using the automated algorithm were then curated by humans ($p < 0.001$; OR 8.56; 95%CI 6.89:10.63), with 51.2 % articles in medical journals curated out of the automated indexed but only 10.9 % in pharmacy practice journals (Table 2).

In total, 59 articles had no MeSH attributed, 43 in medical journals (0.7 %) and 16 in pharmacy journals (0.7 %). The overall number of MeSH attributed per paper was 13.4 (SD 5.0). Table 3 shows the number of MeSH attributed per article in each journal by the three indexing methods. The number of MeSH attributed was different ($p < 0.001$; Cohen's d = 0.941; 95%CI 0.891:0.991) between big-five medical journals (mean 14.6; SD 4.7) and pharmacy practice journals (mean 10.2; SD 4.4). Significant differences in the number of MeSH per article, with moderate effect size, were found in both groups between the different indexing methods (eta-square 0.209; 95%CI 0.191:0.225 for medical journals and 0.409; 95%CI 0.380:0.436 in pharmacy practice journals). Manually indexed articles had a statistically significant higher number of MeSH attributed than automated-indexed articles. However, in both groups the number of MeSH terms attributed when the articles were humanly curated after automated indexing increased (Table 4).

Table 2
Articles indexed by the different indexing methods.

	Manual	Automated and curated	Automated only	Total
	N (%)	N (%)	N (%)	N
Journal				
Big-five medical	4866 (77.8)	710 (11.4)	678 (10.8)	6254
Ann Intern Med	617 (73.9)	42 (5.0)	176 (21.1)	835
BMJ	792 (79.0)	29 (2.9)	182 (18.1)	1003
JAMA	1077 (77.9)	245 (17.7)	61 (4.4)	1383
Lancet	1059 (78.4)	46 (3.4)	246 (18.2)	1351
N Engl J Med	1321 (78.5)	348 (20.7)	13 (0.8)	1682
Pharmacy practice	1124 (50.5)	120 (5.4)	981 (44.1)	2225
Int J Clin Pharm	500 (58.8)	31 (3.6)	320 (37.6)	851
Int J Pharm Pract	178 (53.1)	10 (3.0)	147 (43.9)	335
Res Social Adm Pharm	446 (42.9)	79 (7.6)	514 (49.5)	1039
Origin				
U.S. journals	3461 (70.1)	714 (14.5)	764 (15.5)	4939
Non-U.S. journals	2529 (71.4)	116 (3.3)	895 (25.3)	3540

Also, significant differences existed between journal groups in each indexing method (all p-values <0.001), with medical journals having greater number of MeSH attributed per article in all the three indexing methods, compared to pharmacy practice journals (effect size measures presented in Table 4).

The multivariate analysis of the number of MeSH per article produced a robust prediction model (R-square = 0.361; Durbin-Watson = 1.669; VIF<1.2 for the three covariates). There was a statistical association with journal group ($P < 0.001$; B = 2.495; 95%CI 2.290:2.701) and indexing method ($p < 0.001$; B = 3.044; 95%CI 2.932:3.156) but not with geographic origin (p = 0.126; B = 0.135; 95%IC -0.038:0308).

"Humans" was the most prevalent MeSH term among all the articles, with only 28 (0.3 %) missing this term, which was similar between medical journals (0.4 %) and pharmacy practice journals (0.2 %).

Table 1
Distribution of articles by indexing method and publication year.

Publication year	Big-five medical; N(%)				Clinical & social pharmacy practice; N(%)				Total; N(%)			
	N	Manual	Automated and curated	Automated only	N	Manual	Automated and curated	Automated only	N	Manual	Automated and curated	Automated only
2016	796	776 (97.5)	0	20 (2.5)	171	171 (100)	0	0	967	947 (97.9)	0	20 (2.1)
2017	796	793 (99.6)	0	3 (0.4)	206	206 (100)	0	0	1002	999 (99.7)	0	3 (0.3)
2018	806	805 (99.9)	0	1 (0.1)	259	259 (100)	0	0	1065	1064 (99.9)	0	1 (0.1)
2019	802	801 (99.9)	0	1 (0.1)	310	310 (100)	0	0	1112	1111 (99.9)	0	1 (0.1)
2020	764	760 (99.5)	2 (0.3)	2 (0.3)	306	154 (50.3)	61 (19.9)	91 (29.7)	1070	914 (85.4)	63 (5.9)	93 (8.7)
2021	770	767 (99.6)	0	3 (0.4)	326	24 (7.4)	22 (6.7)	280 (85.9)	1096	791 (72.2)	22 (2)	283 (25.8)
2022	779	164 (21.1)	316 (40.6)	299 (38.4)	335	0	22 (6.6)	313 (93.4)	1114	164 (14.7)	338 (30.3)	612 (54.9)
2023	741	0	392 (52.9)	349 (47.1)	312	0	15 (4.8)	297 (95.2)	1053	0	407 (38.7)	646 (61.3)
Total	6254	4866 (77.8)	710 (11.4)	678 (10.8)	2225	1124 (50.5)	120 (5.4)	981 (44.1)	8479	5990 (70.6)	830 (9.8)	1659 (19.6)

Table 3
Number of Medical Subject Headings (MeSH) assigned per article by journal and indexing method.

Journal	Manual	Automated and curated	Automated only
	N (SD)	N (SD)	N (SD)
Big-five medical			
Ann Intern Med	13.3 (4.5)	10.2 (3.2)	8.9 (2.9)
BMJ	15.0 (4.8)	10.4 (3.4)	9.2 (2.5)
JAMA	15.6 (4.2)	14.1 (3.5)	10.6 (3.3)
Lancet	15.8 (4.4)	10.8 (2.7)	9.6 (2.8)
N Engl J Med	17.0 (3.5)	10.8 (3.0)	8.1 (2.3)
Pharmacy practice			
Int J Clin Pharm	13.9 (4.1)	8.8 (2.6)	7.8 (1.9)
Int J Pharm Pract	12.6 (3.9)	9.2 (3.2)	7.2 (2.0)
Res Social Adm Pharm	12.1 (4.4)	7.3 (2.2)	7.2 (2.1)
Origin			
U.S. journals	15.3 (4.4)	11.7 (3.8)	7.8 (2.7)
Non-U.S. journals	15.0 (4.5)	10.0 (3.0)	8.5 (2.5)

Table 4
Differences in the number of MeSH per article between journal groups and indexing methods.

Indexing method	Big-five medical	Pharmacy practice	p-value[a] (Cohen's d; 95%CI)
	N (SD)	N (SD)	
Manual	15.6 (4.4)	13.0 (4.2)	<0.001 (0.606; 0.541:0.672)
Automated and curated	12.1 (3.5)	7.8 (2.5)	<0.001 (1.250; 1.047:1.452)
Automated only	9.4 (2.8)	7.4 (2.0)	<0.001 (0.838; 0.736:0.940)
p-value[b]	<0.001	<0.001	
eta-squared	0.209	0.409	
95%CI	0.191:0.225	0.380:0.436	

[a] Student's *t*-test.
[b] ANOVA.

Gender-identification MeSH were unevenly distributed, with 64.9 % articles indexed with "Female" and 58.9 % indexed with "Male". For both gender-identification MeSH, a different prevalence was found between journal groups. "Female" MeSH was present in 73.4 % of medical journal articles and only in 41.1 % of pharmacy practice ($p < 0.001$; OR 3.95; 95%IC 3.57:4.37). Similarly, the MeSH "Male" existed in 65.9 % medical journal articles and in 39.1 % pharmacy practice articles ($p < 0.001$; OR 3.02; 95%CI 2.73:3.34).

Of the 8420 articles indexed with any MeSH, 1312 (15.6 %) had at least one of the 31 pharmacy-specific MeSH attributed, with 13 articles (0.2 %) published in medical journals and 1299 articles (58.8 %) in pharmacy practice journals. Articles in pharmacy practice journals were indexed with a mean of 1.1 (SD 1.2) pharmacy-specific MeSH. No difference ($p = 0.714$) in the number of pharmacy-specific MeSH attributed per article was found between the different indexing methods, with a mean of 1.1 (SD 1.0) in manually indexed, 1.1 (SD 1.3) in automated only, and 1.2 (SD 1.3) in the automated plus curated records. A significant difference with small effect size ($p < 0.001$; Cohen's $d = 0.173$; 95%CI 0.089:0.257) in the number of pharmacy-specific MeSH was found between articles published in U.S. journals (mean 1.2; SD 1.2) and non-US journals (1.0; SD 1.1). "Pharmacists" was the most attributed pharmacy-specific MeSH among the articles indexed in pharmacy practice journals with an overall 44.4 % articles, followed by "Community pharmacy services" (22.5 %). Seven of the 31 pharmacy-specific MeSH were never used among the 2225 articles (Table 5). Similar patterns existed among all the pharmacy-specific MeSH across the different indexing methods.

Table 5
Distribution of the pharmacy-specific MeSH among original research articles published in pharmacy practice journals.

	Total, n (%) (n = 2225)	Manual, n (%) (n = 1124)	Automated and curated, n (%) (n = 120)	Automated only, n (%) (n = 981)
Pharmacists	988 (44.4)	509 (45.3)	59 (49.2)	420 (42.8)
Community pharmacy services	501 (22.5)	287 (25.5)	24 (20.0)	190 (19.4)
Pharmacies	275 (12.4)	73 (6.5)	18 (15.0)	184 (18.8)
Pharmacy service, hospital	192 (8.6)	135 (12.0)	11 (9.2)	46 (4.7)
Pharmaceutical services	169 (7.6)	82 (7.3)	8 (6.7)	79 (8.1)
Pharmacy	134 (6.0)	8 (0.7)	16 (13.3)	110 (11.2)
Education, pharmacy	69 (3.1)	43 (3.8)	1 (0.8)	25 (2.5)
Students, pharmacy	59 (2.7)	35 (3.1)	1 (0.8)	23 (2.3)
Education, pharmacy, continuing	15 (0.7)	13 (1.2)	0	2 (0.2)
Pharmacy Research	15 (0.7)	10 (0.9)	1 (0.8)	4 (0.4)
Schools, pharmacy	13 (0.6)	6 (0.5)	0	7 (0.7)
Drug compounding	10 (0.4)	9 (0.8)	0	1 (0.1)
Faculty, Pharmacy	7 (0.3)	4 (0.4)	0	3 (0.3)
Ethics, pharmacy	6 (0.3)	5 (0.4)	0	1 (0.1)
Legislation, pharmacy	5 (0.2)	5 (0.4)	0	0
Clinical pharmacy information systems	3 (0.1)	3 (0.3)	0	0
Insurance, pharmaceutical services	3 (0.1)	3 (0.3)	0	0
Economics, Pharmaceutical	2 (0.1)	1 (0.1)	1 (0.8)	0
Education, pharmacy, graduate	2 (0.1)	2 (0.2)	0	0
Fees, pharmaceutical	2 (0.1)	2 (0.2)	0	0
Pharmaceutical services, online	2 (0.1)	2 (0.2)	0	0
Pharmacy and Therapeutics Committee	2 (0.1)	1 (0.1)	0	1 (0.1)
Societies, pharmaceutical	2 (0.1)	2 (0.2)	0	0
Evidence-Based Pharmacy Practice	1 (0.0)	0	1 (0.8)	0
Behind-the-counter drugs	0	0	0	0
History of Pharmacy	0	0	0	0
Licensure, pharmacy	0	0	0	0
Nuclear Pharmacy	0	0	0	0
Pharmacists' aides	0	0	0	0
Pharmacy administration	0	0	0	0
Practice Patterns, Pharmacists'	0	0	0	0

*List of pharmacy-specific MeSH obtained from the studies doi: 10.2146/ajhp140073; doi: 10.2146/ajhp170046; doi: 10.1016/j.sapharm.2014.11.004

Overall, 1034 articles (12.2 %) contain the word "Pharmacy" in their titles or abstracts, 1001 among pharmacy practice journals (45.0 %) and 33 (0.5 %) in medical journals, being 131 (12.7 %) indexed with the MeSH Pharmacy. Also, 1210 articles contain the word "Pharmacist" in

their titles or abstracts, 1186 (53.3 %) in pharmacy practice journals and 24 (0.4 %) in medical journals, being 951 (78.6 %) indexed with the MeSH "Pharmacists". Failing to allocate the MeSH Pharmacist to an article containing the word "Pharmacist" in the title or abstract was more common among medical journals, with 82.6 % articles missing the MeSH, than in pharmacy practice journals, with only 19.3 % missing the MeSH Pharmacists.

4. Discussion

This study analyzed the MeSH attributed to all the original research articles published between 2016 and 2023 by the big-five medical journals and the three pharmacy practice journals from the Granada Group indexed in MEDLINE. Automated indexing resulted in a reduction of the number of MeSH attributed both to medical and pharmacy journal articles. On articles indexed using the automated method, additional curation by humans increased the amount of MeSH in both journal groups. Pharmacy practice journal articles had significantly fewer MeSH allocated per article than medical journal articles in any indexing method. This lower MeSH presence in pharmacy articles is evident also for generic MeSH such as Humans, Male and Female. Also, the number of MeSH in U.S. based journals was slightly higher than the non-US based journals in both groups, but U.S. based medical journals were more likely than pharmacy practice journals to be curated after the automated indexing process.

Few studies have analyzed the effects of automated indexing in MEDLINE on the number of MeSH allocated per paper. In 2021 Rae et al., when analyzing automation and subheading selection, reported that automated recommendations had high precision but low recall (sensitivity), with some pairs of MeSH terms frequently used, while others were rarely used.[22] In this study, the NLM staff tested two different automated methods against the MTI, showing three times better recall and better consistency for MeSH pair assignment.[22] However, in BioASQ 2022, an annual competition where different automated indexing methods are tested against each other, the NLM's MTI performed below the average (measured in micro-averaged F-measure - MiF) among the systems tested.[23] A recent study conducted by Chen et al., suggested that MTI may not resolve the inconsistencies already observed in manually indexing, thus perpetuating inequities across journals.[16] While our results are in line with these observations, one might expect that differences in the number of MeSH between medical and pharmacy journals would disappear after the automated indexing, which has not happened.

MeSH allocation has been criticized in several health fields.[8,24,25] MeSH allocation by human indexers in the pharmacy practice field has been previously criticized.[12,26–28] The results presented in this study highlight that inaccuracies are being perpetuated in the fully automated indexing method. The bias against pharmacy practice, compared with general medical journals, is obvious in the results of the multivariate analysis, where the U.S. origin lost the significance while only the indexing method and the journal group were associated to a lower number of MeSH. Not only are pharmacy practice articles indexed with less MeSH terms than general medical articles, but they have also fewer generic MeSH terms (e.g., Humans, Male, Female). The lower number of MeSH attributed to pharmacy practice journal articles may reduce their visibility, and subsequent access and citation, which could contribute to lower visibility of pharmacy journals and, more important, the research conducted in this context. Consequently, systematic reviews in the field are less likely to include all relevant outputs, limiting the quality of evidence generation with implications for policy development, professional practice and ultimately health outcomes.

As in previous studies,[27] more than 40 % of the articles published in pharmacy practice journals were not indexed with any of the 31 pharmacy-specific MeSH. The MeSH Pharmacy is an example of a MeSH rarely attributed to pharmacy practice articles. In Minguet et al.,[12] Pharmacy was used in only 2.2 % of the articles published in 10 pharmacy journals, which is similar to the 6.0 % in the present study on pharmacy practice journals. One of the potential reasons for this suboptimal allocation of the MeSH Pharmacy is the limited definition of the MeSH by the NLM: "The practice of compounding and dispensing medicinal preparations". Conversely, the MeSH Pharmacists performed better in previous studies, with 23.6 % articles in Minguet et al.[12] and 23.7 % in Tonin et al.[27] In the present study, more than 44 % articles had this MeSH attributed.

It is important to note that the MTIA only processes the text from articles' titles and abstracts to allocate MeSH terms. The poor standardization in terminology used in abstracts was acknowledged even before automated indexing systems were implemented.[29] It is of the utmost importance that authors and peer reviewers strongly promote the inclusion of MeSH terms as free text words in the abstracts to ensure that the algorithm allocates the accurate MeSH terms, as recommended in the Granada Statements (Recommendations #3 and #4) for pharmacy practice journals (published in 14 journals).[30–43] Of note, several authors have proposed automated methods that use the full text of articles.[44] However, future analyses should focus on the accuracy of the indexing algorithm among articles with appropriate abstracts. In our study, 21 % of articles that included the word Pharmacist in their titles or abstracts were not indexed with this MeSH.

Also, it is worthwhile considering who should allocate the MeSH to index an article. Before 1990, journal abstracts were rarely created by articles' authors. It was only after 1914 that author-generated abstracts became mandatory for all articles in specific journals. Controversy about who should write the abstract existed until 1940s.[45] Today, authors are responsible for preparing the abstract. The importance of ensuring the abstract is sufficiently informative cannot be underestimated as a poor abstract will lead to restricted visibility of the article and similarly, poorly selected MeSH terms will hinder article retrieval from bibliographic databases and lead to limited visibility of the article. There is merit in proposing that authors are responsible for selecting the MeSH terms for their articles.

4.1. Limitations

As in any meta-research exercise, retrieving all the literature from the topic is not guaranteed, especially when inconsistent terminology and procedures are in place. However, to minimize this potential bias we employed comprehensive search strategies to select the original research articles. Also, selected journals included all pharmacy practice journals indexed in MEDLINE, but only a selected group of general medical journals. Results for this second group may not be generalizable to all the medical journals indexed in MEDLINE.

5. Conclusion

Before the implementation of automated indexing methods by the NLM, articles published in pharmacy practice journals were indexed with fewer MeSH terms than those published in general medical journals. This has not changed following the implementation of automated indexing. Almost half of the pharmacy practice articles manually indexed had no pharmacy-specific MeSH terms attributed. This situation has not changed with the implementation of the automated indexing process.

Funding

No external funding.

CRediT authorship contribution statement

Fernando Fernandez-Llimos: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Luciana G. Negrão:** Data curation, Formal analysis, Writing –

original draft, Writing – review & editing. **Christine Bond:** Conceptualization, Methodology, Writing – review & editing. **Derek Stewart:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Fernando Fernandez-Llimos is associate editor of Revista Brasileira de Farmácia Hospitalar e Serviços de Saúde. Christine Bond is editor-in-chief of International Journal of Pharmacy Practice. Derek Stewart is editor-in-chief of International Journal of Clinical Pharmacy. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sapharm.2024.06.003.

References

1. National Library of Medicine. Medical subject headings: preface. https://www.nlm.nih.gov/mesh/intro_preface.html. Accessed June 1, 2024.
2. National Library of Medicine. Module 1: foundations of MeSH in MEDLINE®. https://www.nlm.nih.gov/oet/ed/pubmed/mesh/mod01/index.html. Accessed June 1, 2024.
3. Jenuwine ES, Floyd JA. Comparison of Medical Subject Headings and text-word searches in MEDLINE to retrieve studies on sleep in healthy individuals. *J Med Libr Assoc*. 2004;92(3):349–353.
4. Golder S, Loke YK. Sensitivity and precision of adverse effects search filters in MEDLINE and EMBASE: a case study of fractures with thiazolidinediones. *Health Inf Libr J*. 2012;29(1):28–38. https://doi.org/10.1111/j.1471-1842.2011.00972.x.
5. Shultz M. Comparing test searches in PubMed and google scholar. *J Med Libr Assoc*. 2007;95(4):442–445. https://doi.org/10.3163/1536-5050.95.4.442.
6. Bramer WM, de Jonge GB, Rethlefsen ML, et al. A systematic approach to searching: an efficient and complete method to develop literature searches. *J Med Libr Assoc*. 2018;106(4):531–541. https://doi.org/10.5195/jmla.2018.283.
7. Minguet F, Van Den Boogerd L, Salgado TM, et al. Characterization of the medical subject headings thesaurus for pharmacy. *Am J Health Syst Pharm*. 2014;71(22):1965–1972. https://doi.org/10.2146/ajhp140073.
8. Portaluppi F. Consistency and accuracy of the Medical Subject Headings thesaurus for electronic indexing and retrieval of chronobiologic references. *Chronobiol Int*. 2007;24(6):1213–1229. https://doi.org/10.1080/07420520701791570.
9. Salgado TM, Fernandez-Llimos F. Missing pharmacy-specific medical subject headings (MeSH) terms: problems and solutions. *Res Soc Adm Pharm*. 2019;15(9):1189–1190. https://doi.org/10.1016/j.sapharm.2019.04.008.
10. Fernandez-Llimos F, Minguet F, Salgado TM. New pharmacy-specific medical subject headings included in the 2017 database. *Am J Health Syst Pharm*. 2017;74(15):1128–1129. https://doi.org/10.2146/ajhp170046.
11. Richter RR, Austin TM. Using MeSH (medical subject headings) to enhance PubMed search strategies for evidence-based practice in physical therapy. *Phys Ther*. 2012;92(1):124–132. https://doi.org/10.2522/ptj.20100178.
12. Minguet F, Salgado TM, van den Boogerd L, et al. Quality of pharmacy-specific Medical Subject Headings (MeSH) assignment in pharmacy journals indexed in MEDLINE. *Res Soc Adm Pharm*. 2015;11(5):686–695. https://doi.org/10.1016/j.sapharm.2014.11.004.
13. Rodriguez RW. Comparison of indexing times among articles from medical, nursing, and pharmacy journals. *Am J Health Syst Pharm*. 2016;73(8):569–575. https://doi.org/10.2146/ajhp150319.
14. Mork J, Aronson A, Demner-Fushman D. 12 years on - is the NLM medical text indexer still useful and relevant? *J Biomed Semant*. 2017;8(1):8. https://doi.org/10.1186/s13326-017-0113-5.
15. National Library of Medicine. Frequently asked questions about indexing for MEDLINE. Available from:: https://www.nlm.nih.gov/bsd/indexfaq.html. (cited 1-June-2024).
16. Chen E, Bullard J, Giustini D. Automated indexing using NLM's Medical Text Indexer (MTI) compared to human indexing in Medline: a pilot study. *J Med Libr Assoc*. 2023;111(3):684–694. https://doi.org/10.5195/jmla.2023.1588.
17. Allotey P, Allotey-Reidpath C, Reidpath DD. Gender bias in clinical case reports: a cross-sectional study of the "big five" medical journals. *PLoS One*. 2017;12(5), e0177386. https://doi.org/10.1371/journal.pone.0177386.
18. Farrow L, Gardner WT, Ablett AD, et al. A review of trauma and orthopaedic randomised clinical trials published in high-impact general medical journals. *Eur J Orthop Surg Traumatol*. 2022;32(8):1469–1479. https://doi.org/10.1007/s00590-021-03137-3.
19. Pressat-Laffouilhere T, Jouffroy R, Leguillou A, et al. Variable selection methods were poorly reported but rarely misused in major medical journals: literature review. *J Clin Epidemiol*. 2021;139:12–19. https://doi.org/10.1016/j.jclinepi.2021.07.006.
20. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. *Am Statistician*. 2016;70(2):129–133. https://doi.org/10.1080/00031305.2016.1154108.
21. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum; 1988.
22. Rae AR, Pritchard DO, Mork JG, et al. Automatic MeSH indexing: revisiting the subheading attachment problem. *AMIA Annu Symp Proc*. 2020;2020:1031–1040.
23. Krithara A, Mork JG, Nentidis A, et al. The road from manual to automatic semantic indexing of biomedical literature: a 10 years journey. *Front Res Metr Anal*. 2023;8, 1250930. https://doi.org/10.3389/frma.2023.1250930.
24. Murphy LS, Reinsch S, Najm WI, et al. Searching biomedical databases on complementary medicine: the use of controlled vocabulary among authors, indexers and investigators. *BMC Compl Alternative Med*. 2003;3:3. https://doi.org/10.1186/1472-6882-3-3.
25. Wieland S, Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *J Clin Epidemiol*. 2005;58(6):560–567. https://doi.org/10.1016/j.jclinepi.2004.11.018.
26. Minguet F, Salgado TM, Santopadre C, et al. Redefining the pharmacology and pharmacy subject category in the journal citation reports using medical subject headings (MeSH). *Int J Clin Pharm*. 2017;39(5):989–997. https://doi.org/10.1007/s11096-017-0527-2.
27. Tonin FS, Gmunder V, Bonetti AF, et al. Use of 'pharmaceutical services' medical subject headings (MeSH) in articles assessing pharmacists' interventions. *Explor Res Clin Soc Pharm*. 2022;7, 100172. https://doi.org/10.1016/j.rcsop.2022.100172.
28. Rotta I, Souza TT, Salgado TM, et al. Characterization of published randomized controlled trials assessing clinical pharmacy services around the world. *Res Soc Adm Pharm*. 2017;13(1):201–208. https://doi.org/10.1016/j.sapharm.2016.01.003.
29. Nedobity W. The relevance of terminologies for automatic abstracting. *J Inf Sci*. 2016;4(4):161–165. https://doi.org/10.1177/016555158200400405.
30. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada statements. *Res Soc Adm Pharm*. 2023;19(5):830–835. https://doi.org/10.1016/j.sapharm.2023.01.007.
31. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada Statements. *Int J Clin Pharm*. 2023;45(2):285–292. https://doi.org/10.1007/s11096-023-01550-8.
32. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada Statements. *Explor Res Clin Soc Pharm*. 2023;9, 100229. https://doi.org/10.1016/j.rcsop.2023.100229.
33. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada statements. *Int J Pharm Pract*. 2023;31(2):119–125. https://doi.org/10.1093/ijpp/riad017.
34. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada Statements. *Eur J Hosp Pharm*. 2023. https://doi.org/10.1136/ejhpharm-2023-003748.
35. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada statements. *J Pharm Policy Pract*. 2023;16(1):43. https://doi.org/10.1186/s40545-023-00527-2.
36. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada Statements. *Farm Hosp*. 2023;47(3):133–138. https://doi.org/10.1016/j.farma.2023.01.006.
37. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the entire paradigms of clinical and social pharmacy practice research: the Granada statements. *Ars Pharm*. 2023;64(2):161–172. https://doi.org/10.30827/ars.v64i2.27182.
38. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada statements. *Pharm Educ*. 2023;23(1):109–117. https://doi.org/10.46542/pe.2023.231.109117.
39. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada statements. *Revista Brasileira de Farmácia Hospitalar e Serviços de Saúde*. 2023;14(1). https://doi.org/10.30968/rbfhss.2023.141.0913.
40. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the entire paradigms of clinical and social pharmacy practice research: the Granada statements. *J Pharm Clin*. 2023;42(2):53–60. https://doi.org/10.1684/jpc.2023.0523.
41. Fernandez-Llimos FF. Mejorando la calidad de las publicaciones y avanzando en todos los paradigmas de la investigación de la farmacia asistencial, clínica y social: las declaraciones de Granada. *Farmacéuticos Comunitarios*. 2023;15(4):31–38. https://doi.org/10.33620/fc.2173-9218 (2023).23.
42. Fernandez-Llimos F, Desselle S, Stewart D, et al. Improving the quality of publications in and advancing the paradigms of clinical and social pharmacy practice research: the Granada statements. *J Pharmaceut Health Serv Res*. 2023. https://doi.org/10.1093/jphsr/rmad039.
43. Fernandez-Llimos F, Desselle S, Stewart D, et al. Mejora de la Calidad de las Publicaciones y Avance en los Paradigmas de la Investigación en Farmacia Práctica

Clínica y Social: La Declaración de Granada. *Pharm Care Espana*. 2023;25(2):65–99. https://doi.org/10.60103/phc.v25i2.816.
44. Dai S, You R, Lu Z, et al. FullMeSH: improving large-scale MeSH indexing with full text. *Bioinformatics*. 2020;36(5):1533–1541. https://doi.org/10.1093/bioinformatics/btz756.
45. Royal Society. Where did the practice of 'abstracts' come from? Available from:: https://arts.st-andrews.ac.uk/philosophicaltransactions/where-did-the-practice-of-abstracts-come-from/. (cited 1-June-2024).