MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190217

Examining Reproducibility of Literature Search in Meta-Analysis

Fan Li[a], Pei-Yin Yang[b], Hong Kang[b], Xiaoqiu Chen[c], Chao Shang[d], Yang Gong[b]

[a] *School of Medical Informatics, China Medical University, Shenyang, Liaoning Province, China,*
[b] *School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, Texas, United States,*
[c] *Department of ENT, Shengjing Hospital of China Medical University, Shenyang, Liaoning Province, China,*
[d] *Network Center of China Medical University, Shenyang, Liaoning Province, China*

Abstract

Meta-analysis, a systematic retrieval from literature databases is an essential and prevailing method for combining data from multiple studies. Unfortunately, few studies have examined its rigor, which affects its reproducibility of results. We identified 22 meta-analyses on cervical cancer in PubMed for examining the parameters defined by PRISMA, relating to the rigor of literature retrieval. We found that 16 literature databases were used, and EMBASE was a leading resource, accounting for the highest frequency (81.82%). About half (45.45%) of the meta-analyses presented a complete, reproducible search strategy for at least one database. The ratio of included to retrieved articles after redundancy removal was only 6.58%, indicating low precision due to unclear or unreported processes. Our work serves as an initial step to examine the planning and execution of meta-analysis. Future efforts need to enhance reliability on literature retrieval in meta-analysis and compliance to PRISMA.

Keywords:

Meta-Analysis; Information Storage and Retrieval; Reproducibility of Results

Introduction

It has been an essential yet challenging task for clinicians to stay aprised on new knowledge from primary research papers while practicing in clinical settings [1]. Meta-analysis, the highest hierarchy of evidence regarding intervention questions [2], is defined as "the statistical synthesis of individual patient data from varying primary studies, leading to a quantitative summary of the pooled results" [3]. Meta-analyses assist clinical decision making, guide evidence-based medicine (EBM) and clinical practice, and serve as "the policy foundation for evidence-based practice guidelines, economic evaluations and future research agenda" [4].

The quality of the above-mentioned evidence is important. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) was developed for the reporting of systematic reviews and meta-analyses. PRISMA is a 27-item checklist and four-phase flow diagram that can examine whether the systematic reviews and meta-analyses are able to identify, appraise, and summarize research in an objective fashion so that clinicians would know whether the information is reliable for decision-making; however, most reported meta-analyses do not include such a quality assessment tool or do not report the method of assessment [5]. AMSTAR, an 11-item checklist, has been employed to assess the quality of meta-analysis and systematic review. Prior studies on quality assessments have identified that the Cochrane library was of a higher quality than others. Investigations of study selection bias and data extractions returned unclear results [6].

Systematic reviews require an unbiased and a reproducible search of data resources to identify as many relevant studies as possible. Reliability is the quality or state of being reliable, the extent to which an experiment, test, or measuring procedure yields the same results on repeated trials. In the field of academic research, reproducibility refers to that for the same research problem, independent research by other researchers can use the scheme provided by the authors to reproduce the experimental results [1]. PRISMA described Section Method "Search" as "Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated" and provided a guideline for researchers to examine the reproducibility of systematic reviews and meta-analyses [7]; however, researchers have conducted few meticulous studies on examining the reproducibility of meta-analyses.

This paper aims to examine the reproducibility of search strategy reported in meta-analyses. We chose cervical cancer, because providers can often detect it early, and sometimes prevent it entirely, by having regular Pap tests. When found early, cervical cancer is one of the most successfully treatable cancers. In addition, cervical cancer is the second most common cancer among females worldwide, with 80% of the cases occurring in sub-Saharan Africa, Central America, and South-Central Asia [8]. We aim to examine the reproducibility of search strategies reported by cervical-cancer related meta-analyses and reveal the opportunities and direction for clinical informaticians toward an enhanced rigor of meta-analyses.

Methods

First, we performed a systematic literature search in PubMed database to identify meta-analyses focusing on cervical cancer published from January 1st, 2013 to October 3rd, 2018. The search strategy utilized search terms as follows: (cervical cancer) AND (meta-analysis[PT] OR meta-analysis[tiab])

Titles and abstracts of the retrieved articles were screened manually and independently by two authors (FL and PY) using the eligibility criteria. The inclusion criteria were: 1) studies or systematic reviews with meta-analysis, data synthesis or quantitative overview; 2) studies focusing on cervical cancer. The exclusion criteria were: 1) comments or corrections for articles of meta-analysis; 2) narrative reviews or meta-analyses focusing on other cancers, which are different than cervical cancer, such as head and neck cancer. Figure 1 illustrates the overall strategy.

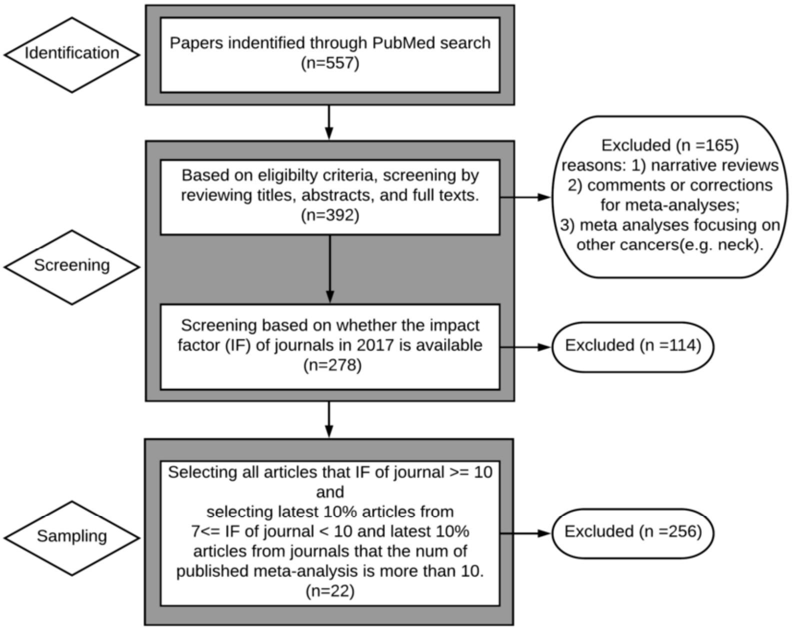

Figure 1- Flow Chart of Literature Search and Inclusion

Second, a Bibliographic Item Co-Occurrence Matrix Builder (BICOMB) analyzed the journal distribution of these articles [9]. The 2017 Journal Citation Report provided the impact factors (IF) of the journals. Next, we applied a stratified sampling strategy to the retrieved meta-analyses for further data extraction and evaluation: 1) all articles published in the journals with an IF higher than 10, 2) 10% of latest published articles in the journals with an IF from 7 to 10, and 3) 10% of latest published articles in the journals containing more than 10 articles, i.e. two gynecologic journals (*Gynecologic Oncology* and *Archives of Gynecology & Obstetrics*) and a comprehensive journal (*Plos One*).

Following the checklist of PRISMA, features of evaluation in this study were created, including journal name, publication year, electronic databases used, presence of full search strategy for all databases or at least one database, the number of retrieved articles before and after the removal of duplicates, and the number of included articles. Two authors (FL and PY) independently extracted the features from the full papers according to the checklist. Group discussions were held to resolve discrepancies involving additional authors when necessary. For each meta-analysis, we calculated the percentage of the number of included articles to the number of retrieved articles (before and after removal of duplicates).

Results

Search Results and Characteristics of Meta-Analyses

557 articles were retrieved from PubMed by applying the search strategy, among which 392 articles were included by applying the eligibility criteria. The 392 articles were published in a total of 157 journals, among which 123 (78.3%) journals, carrying 278 (70.9%) articles, possessed impact factors in the 2017 Journal Citation Report. A journal list with ranked IF in descending order and published in 2017 was created. Figure 2 illustrates the distribution of IF of journals and number of relevant meta-analyses on cervical cancer. The application of the stratified strategy resulted in 22 articles, shown in Table 1.

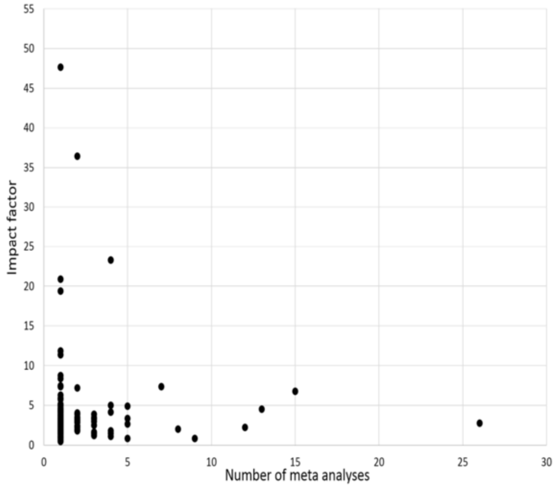

Figure 2- Distribution of Impact Factors of Journals and Number of Relevant Meta-Analyses on Cervical Cancer

Note: The journals (IF=2 to 7） contain more than 10 meta-analyses. Journals tend to publish meta-analyses on cervical cancer with IF < 10.

Table 1- Characteristics of Meta-Analyses Selected for Assessment

Author Year PMID	Journal	Search strategy§		# retrieved articles€		# included articles	% of included/retrieved articles£	
		1§	2§	1€	2€		1£	2£
Melnikow 2018 30140883	JAMA	Y*	Y	5232	2972	62	1.19	2.09
Arbyn 2017 29126708	Lancet. Oncology	N$	Y	N	N	93	N	N
Arbyn 2014 24433684	Lancet. Oncology	Y	Y	?л	?(884£)	97	N	N
Tainio 2018 29487049	BMJ	Y	Y	?(6275)	?	36	N	N
Fokom-Domgue 2015 26142020	BMJ	Y	Y	?	?(1049)	15	N	N
Kyrgiou 2014 25352501	BMJ	Y	Y	?	?(1697)	15	N	N
Kyrgiou 2016 27469988	BMJ	Y	Y	?	?(3021)	71	N	N
Arbyn 2017 27842420	Annals of Internal Medicine	N	Y	?	?(899)	24	N	N
Siristatidis 2013 23255514	Human Reproduction Update	Y	Y	7785	7785	9	0.12	0.12
Kelly 2018 29107561	Lancet HIV	Y	Y	605	407	16	2.64	3.93
Zard 2014 24657969	Autoimmunity Reviews	N	N	?	?(235)	7	2.98	N
Fisher 2013 23620381	International Journal of Epidemiology	N	N	1108	699	29	2.62	4.83
Hammer 2016 26661889	International Journal of Cancer	N	N	721	644	15	2.08	2.33
Li 2014 24308856	Alimentary Pharmacology & Therapeutics	N	N	565	393	63	11.15	16.03
Verdoodt 2015 26296294	European Journal of Cancer	N	N	376	252	16	4.26	6.35
de Lima 2018 29021084	Gynecologic Oncology	N	N	396	N	25	6.31	N
Charakorn 2018 29606483	Gynecologic Oncology	N	N	1797	1605	61	3.39	3.80
Zhang 2018 29641578	Plos One	N	N	1715	1342	20	1.17	1.49
Jin 2018 29554090	Plos One	N	N	2614	2588	N	N	N
Zhou 2017 29227998	Plos One	N	N	709	707	13	1.83	1.84
Ye 2018 29520664	Archives of Gynecology & Obstetrics	N	N	319	225	22	6.90	9.78
Feng 2018 29876746	Archives of Gynecology & Obstetrics	N	N	79	69	8	10.13	11.59

Notes: §: Search strategy 1 represents whether the full electronic search strategy for all databases is available. Search strategy 2 represents whether the full electronic search strategy for at least one database is available.

€: The # retrieved articles 1 is the number of retrieved articles before removal of duplicates, and 2 is that of those after removal of duplicates.

£: For % of included/retrieved articles, the first column represents the percentage of number of included articles to that of retrieved articles before removal of duplicates, and the second column represents the percentage of number of included articles to that of retrieved articles after removal of duplicates.

**: Y: available.; $: N: not available.; л: ?: not clear. ; £: The number in brackets is a rough estimate by the context.*

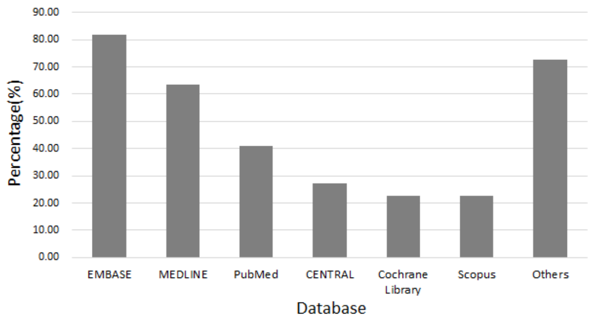

Figure 3- Literature Databases Searched in the Meta-Analyses

Notes: MEDLINE includes PubMed MEDLINE and OVID MEDLINE. Cochrane databases are part of Cochrane Library. Other resources include CINAHL, CNKI, ISI Web of Science, PsycINFO, Wanfang, CBM, Cochrane Database of Systematic Reviews, EBSCO, Google Scholar, ISI Proceedings, and PubMed/MEDLINE with a frequency ≤ 2.

Literature Databases Used in the Meta-Analyses

Seventeen literature databases (Involvement frequencies: Maximum = 9; Minimum = 2; Average = 4) were identified in the 22 meta-analyses. As illustrated in Figure 3, the top three databases involved in the meta-analyses were EMBASE (18 meta-analyses, 81.82%), MEDLINE (including PubMed MEDLINE and OVID MEDLINE, 14 meta-analyses, 63.64%), and PubMed (9 meta-analyses, 40.91%). Nine of 14 meta-analyses in which MEDLINE was searched (64.29%) did not specify the platform of MEDLINE .

Inclusion of One Full Search Strategy

Ten (45.45%) meta-analyses contained a full search strategy including any limits used for at least one database, among which eight (36.36%) meta-analyses reported the full search strategy for all databases. In the high impact journals (IF≥10), all meta-analyses presented a full search strategy for at least one database, while eight (80%) meta-analyses reported the full search strategy for all databases; however, among the low impact journals (IF<10), no meta-analyses presented a full search strategy even for at least one database.

Comparing the Number of Included and Retrieved Articles

The number of retrieved articles from each information resource was not always clearly included, shown in Table 1, resulting in the failures in determining whether or not the total amount of retrieved articles was subject to removal of duplicates. Twelve meta-analyses reported the number of retrieved articles before and after removal of duplicates and the number of included articles, among which the maximum number of retrieved articles after duplicate removal was 2,972, and its corresponding number of included articles was 62. Table 1 also presents the percentage of the number of included articles to the number of retrieved articles after duplicate removal. The percentage ranged from 1.49% to 16.03% (mean= 6.58%).

Discussion

Confusion on Literature Databases as Information Source

Based on the item regarding information sources in PRISMA 2009, it is necessary to describe all information sources (e.g., databases with dates of coverage and author contacts). Although we could provide information sources for the 22 meta-analyses we evaluated, choosing source information is frequently unclear due to the description of the names of databases and the relationship among databases. For example, PubMed and MEDLINE should not be used in the same meta-analysis, since MEDLINE is in fact a subset of PubMed. In addition, searching results may vary in different MEDLINE databases (e.g., PubMed, Ovid MEDLINE) [10]. More than one half of meta-analyses did not report the platform they used to search MEDLINE, which becomes a barrier in reproducing the results or repeating the search strategy.

The number of searched literature databases under the topic of cervical cancer varied between two and nine. The reasons of including multiple databases are extremely unclear and somewhat confusing. For example, databases like Google Scholar and ISI Proceedings are used in addition to specialized professional databases, such as PubMed and EMBASE. It is rarely reported how many additional hits were introduced by such an unclear inclusion of databases. These problems remain for future exploration, and further investigations should clarify the rules on use of databases in meta-analyses.

Poor Compliance on Inclusion of One Full Search Strategy

To establish a reasonable and detailed search strategy for each database and ensure the quality and reliability of meta-analysis, it is essential to enhance recall and precision of evidence. According to the PRISMA criteria, the search strategy should be repeatable for at least one database. Our findings indicate that only the strategies published in the high IF journals tend to be repeatable. It is necessary for meta-analysis to include at least one repeatable full search strategy for at least one database or for every database included in the search.

Lower Precision of Evidence Retrieval

Based on the data from 11 meta-analyses, we could obtain the ratio of the number of included articles to the number of retrieved articles after removal of duplicates, ranging from 1.49% to 16.03% (mean 6.58%). This percentage represents the precision of literature retrieval on meta-analyses. We found that authors usually developed a search strategy for higher recall so that the precision is much lower than general literature retrieval. The heavy workload of article screening may be subjective and prone to errors. This problem could be resolved by establishing a search strategy with clarified keyword limitations, which may reduce recall as a compromise. Therefore, systematic filters of potential eligible articles should be developed.

Lack of Reproducibility

The number of retrieved articles was not clearly presented for each information resource, which could be another barrier to repeat a meta-analysis. Table 1 shows that only three articles (13.64%) presented a full search strategy for all databases as well as the number of retrieved articles before and after removal of duplicates. The absence of full search strategy for all databases in large amounts of meta-analyses weakens the possibility of repeating them. Therefore, one full search strategy for at least one database in PRISMA is inadequate in meta-analysis. A full search strategy and number of article hits for each database should be a fundamental requirement of meta-analysis.

Recommendations

Based on the findings above, we suggest the following recommendations in order to make a meta-analysis repeatable at the literature retrieval stage. First, we propose to enhance the PRISMA criteria to present full electronic search strategies for all databases rather than for at least one. Researchers should strictly obey the item on presentation of detailed operable electronic search strategy according to PRISMA. Second, PRISMA should require that meta-analyses report the number of literature yielded from each database. Third, researchers who will participate in meta-analysis should be trained with the knowledge of information resource retrieval to control various problems that may arise from the literature search process and to ensure the repeatability of the results. In any case, reviewers and editors of journals should rigidly control the quality of literature retrieval for meta-analyses according to PRISMA.

Limitations and Future Work

Only 22 articles were included for the assessment on quality of literature search based upon a stratified sampling of 392 meta-analyses on cervical cancer. The limited sample size may have limited generalizability of the results. Nonetheless, the conclusion may be specific in cervical cancer and hold limited scalability to other gynecologic cancers. In the future, we will evaluate all the 392 meta-analyses by the AMSTAR checklist and perform a univariate and multivariate statistical analysis to investigate the factors influencing the quality of literature retrieval of meta-analyses.

Conclusions

The assessment of meta-analyses in regards to literature retrieval revealed poor performances in reporting quality of retrieval strategy and a low compliance of PRISMA. Overall, the reporting quality of literature retrieval in the meta-analysis on cervical cancer needs to be improved. Specifically, a clear justification is needed in describing the selected databases. PRISMA requires to include a complete search strategy toward reproducible results. The number of article hits retrieved from each and every database should be reported individually instead of a combined number which is a barrier for reproducibility of meta-analysis.

Acknowledgements

This work was supported by the Science and Technology Research Program of Department of Education of Liaoning Province, PR China (Grant No. LK201659). The study was also supported by UTHealth Innovation for Cancer Prevention Research Training Program Post-Doctoral Fellowship (RP160015).

References

[1] M.W. Russo, How to Review a Meta-analysis, *Gastroenterol Hepatol (N Y)* **3** (2007), 637-42.

[2] H. Szajewska, Evidence-Based Medicine and Clinical Research: Both Are Needed, Neither Is Perfect, *Ann Nutr Metab* **72 Suppl 3** (2018), 13-23.

[3] K. O'Rourke, An historical perspective on meta-analysis: Dealing quantitatively with varying study results, *J R Soc Med* **100** (2007), 579-82.

[4] D. Moher, D.J. Cook, S. Eastwood, I. Olkin, D. Rennie, D.F. Stroup, Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. Quality of Reporting of Meta-analyses, *Lancet* **354** (1999), 1896-900.

[5] S. Sanderson, I.D. Tatt, J.P. Higgins, Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: A systematic review and annotated bibliography, *Int J Epidemiol* **36** (2007), 666-76.

[6] J. Dong, G. Teng, T. Wei, W. Gao, H. Wang, Methodological Quality Assessment of Meta-Analyses and Systematic Reviews of Probiotics in Inflammatory Bowel Disease and Pouchitis, *PLoS One* **11** (2016), 1-11, doi:10.1371/journal.pone.0168785.

[7] A. Liberati, D.G. Altman, J. Tetzlaff, C. Mulrow, P.C. Gotzsche, J.P. Ioannidis, M. Clarke, P.J. Devereaux, J. Kleijnen, D. Moher, The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration, *PLoS Med* **6** (2009), doi: 10.1136/bmj.b2700.

[8] C.F.C.C. Collaboration, Reducing uncertainties about the effects of chemoradiotherapy for cervical cancer: A systematic review and meta-analysis of individual patient data from 18 randomized trials, *J Clin Oncol* **26** (2008), 5802-12.

[9] L. Cui, W. Liu, L. Yan, H. Zhang, Y.F. Hou, Y.N. Huang, H. Zhang, Development of a text mining system based on the co-occurrence of bibliographic items in literature, *New Technology of Library and Information Service* (2008), 70-75.

[10] A. Wanner, N. Baumann, Design and implementation of a tool for conversion of search strategies between PubMed and Ovid MEDLINE, *Res Synth Methods* (2018).

Address for correspondence

Yang Gong, MD, PhD
7000 Fannin St Suite 600, Houston, TX 77030, USA
Yang.Gong@uth.tmc.edu
Phone: +1 713 500 3547, Fax: +1 713 500 3907.