Literature Search for Scientific Processes in Medical Devices: Challenges, Errors, and Mitigation Strategies

Ashish Indani*, Srinivas Reddy Boreddy, Nadeem Ansari

TATA Consultancy Services, Mumbai, India

Abstract **-** Literature search is a commonly used strategy or a method of collecting evidence on a given research question, specifically for a Clinical Evaluation Report. A precise literature search not only provides accurate evidence but also saves time and efforts during collation of such data. However, unless implemented correctly, literature search can be misleading, time-consuming, or useless. Focusing the literature search on a precise topic and obtaining relevant evidence in a stipulated time requires high skill levels. Despite several guidance documents and papers, the process of literature search has various types of errors. These are errors of inclusion, exclusion, inclusive exclusion, exclusive inclusion, and limited relevance (exclusive exclusions). In order to obtain optimal outcomes in a literature search, the analysis of these errors is important. These errors pertain to the volume of evidence, relevance of the data, tone of evidence, and its value to the research topic. Analyzing these challenges and devising an accurate strategy to overcome these errors would certainly improve literature search outcomes. Combinations and permutations of these challenges (volume and relevance) present various practical challenges, namely, too high data; too low data; high volume, low relevance data; low volume, low relevance data; high value, high relevance data but repetitive outcomes (monotonous); and high value, high relevance data, but missing trends and threads. In this article, we discuss the above-mentioned errors and challenges and mitigation strategies along with literature search automation.

Index Terms- Literature search, challenges, errors, mitigation strategies

I. Background and Problem Statement

Clinical evaluation report (CER) is a widely accepted document for the approval of medical devices by various regulatory bodies, and literature search plays a pivotal role in drafting a CER. In fact, recent modifications of the MedDEV 2.7.1 guidelines (June 2016) focused more on the role of literature search as a tool for medical device clinical evaluation and on establishing state-of-the-art treatment of diseases in which the device is used. Hence, in the present article, we discussed errors and challenges in literature search while identifying and appraising evidence (references) and strategies to mitigate them.

Clinical evaluation of medical devices is a process similar to synthesizing a systematic review or a research article. Data documentation and publication are the most important aspects of any research,[1] regardless of whether the outcomes of the research are accepted or challenged. When a documented research is cited as a reference, it assures readers of the quality of the research and presents previous knowledge on the topic under investigation. A reference is important in providing information on the past efforts, methods, and strategies used; method or strategy alternatives; intellectual property information; and the current status of the selected research topic. Addition of contemporary and historical research data confirms the validity and feasibility of the research project and also reduces the chances of failure, thereby saving cost and resources.[i] For a medical device industry, such addition helps in reducing failures and thus, reduces the cost of development.

Literature search is a commonly used strategy or a method of collecting evidence on a given research question.[ii] A precise literature search not only provides accurate evidence but also saves time and efforts required in collecting evidence. However, unless implemented correctly, literature search can be misleading, time consuming, or useless. Focusing literature search on a precise topic and obtaining relevant evidence within a stipulated time often demands high skill levels. Despite several guidance documents and papers on its methodology, the literature search process still has various errors. In order to obtain optimum outcomes with literature search, the analysis of these errors is a high priority study. These errors are related to the volume of evidence, relevance of the data, tone of evidence, and its value to the research topic. Analyzing these challenges and devising an accurate strategy to overcome these would certainly improvise literature search outcomes.

The present article defines errors due to the incorrect use of key elements during literature search, such as keywords, database, and Boolean logic. We briefly explain how these errors may manifest into either inclusion of noise or exclusion of relevant information. We also explain how these errors can lead to challenges in literature search. Although prima facie, this white paper is focused on literature search for CER and medical devices, it is generally applicable to all types of literature search.

Errors and challenges in literature search

Errors and challenges in literature search are the main reasons why outcomes deviate. Errors have their origin in incorrect use of primary attributes of literature search, viz., keywords, Boolean, and database. The attributes of these errors are stated below:

a. Errors in setting eligibility criteria (type of literature and databases)
b. Errors in selecting keywords and Boolean logics
c. Errors in setting up the search phrases for database

Often, these attributes concur and can form five types of errors:

1. Error of inclusion
2. Error of exclusion
3. Error of inclusive exclusion

4. Error of exclusive inclusion
5. Error of limited relevance (error of exclusive exclusions)

Error 1 – Error of inclusion

The error of inclusion means including more information than required (Figure 1). This error may occur in case of inadequate exclusion criteria, use of broad keywords, non-specific databases, or Boolean logic misuse. In addition, use of improperly formed long search strings, such as several synonyms joined by the "OR" Boolean logic, may result in the error of inclusion. An error of inclusion may occur even in a topic-specific database if a search phrase contains broad keywords, improper Booleans, or many words with truncations or wildcards.

When an error of inclusion occurs, the overall completion of the task may be delayed because a longtime is required to appraise the literature, while keeping topic relevancy as a prime objective. Furthermore, an error of inclusion leads to high volumes of low value literature, as it returns many irrelevant results and often fetches duplicates. Therefore, the effort required in a literature search is disproportionate to the actual usable data obtained and the time taken in the process. However, in case of rare diseases or treatments, the error of exclusion is preferred over the error of exclusion because the target literature remains in the literature pool.

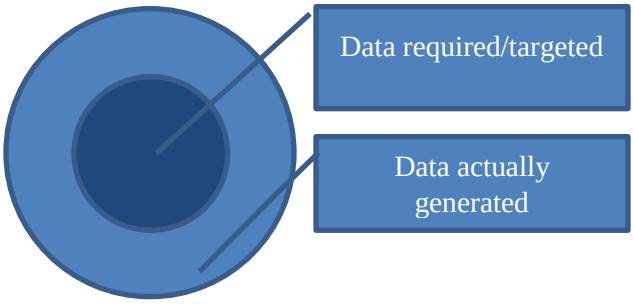


Figure 1. Error of inclusion

Error 2 – Error of exclusion

In an error of exclusion, relevant data are not included in the search because of an extreme exclusion-driven strategy. This error occurs in three scenarios:

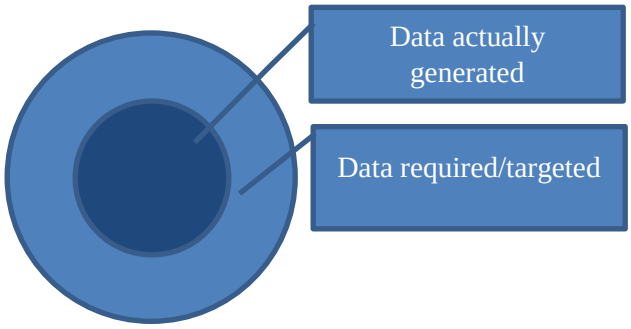
1. When eligibility criteria fails to set correct inclusions
2. Stringent or too specific keywords
3. Use of excluding Boolean logic "AND" or "NAND" to form search strings

Exclusions in the literature can by influenced by literature type, keywords, synonyms, similar words, or phrases, etc. When an error of exclusion occurs, relevant and useful literature may be excluded (Figure 2) from the literature. This error can occur even during appraisal owing to stringent appraisal criteria. An error of exclusion may lead to specific literature outputs that often lead to less information and thus may cause bias.

Error 3 – Error of inclusive exclusions

Errors of inclusive exclusions are due to bias by literature search professionals. Typically, the protocol in this type of error has an appropriate set of exclusion and inclusion criteria; however, because of selection bias, the key terms used in the search may be linked to a specific or desired (leading) outcome, and hence, monotonous data may be returned. MeSH terms or controlled vocabulary are highly recommended for key terms as these words fetch related information. The bias can range from inclusion of only positive results of a particular treatment for a particular disease to the inclusion of only specific brand names and event types. Owing to the use of several synonyms and inclusive Boolean logic such as OR, the search appears to be inclusive or correct search. However, bias in selection and combinations of keywords (inclusive their synonyms and wild cards) leads to errors, and returns one-sided outputs and relevant data may be missed (Figure 3). In some cases, the error of inclusive exclusion also occurs when exclusion criteria are not set to exclude sponsored or promotional literature. This bias tends to shift the data trends in favor of the sponsor. However, some of the bias can be recognized and addressed in the appraisal process.

For example, the search phrase, "Diabetes Mellitus" AND ("Pioglitazone" AND "Normoglycemia") or "Diabetes", AND ("Pioglitazone" AND "Normal blood sugar") or Pioglitazone AND (Normal Blood Sugar OR Normoglycemia OR Normal HbA1c OR HBA1c Less than 7, OR Blood sugar Less than 140), etc., is less likely to generate literature data on adverse events



related to Pioglitazone. Mostly, metabolic adverse events such as hypoglycemia caused by the drug can be anticipated in this search outcome.

Figure 2. Error of exclusion

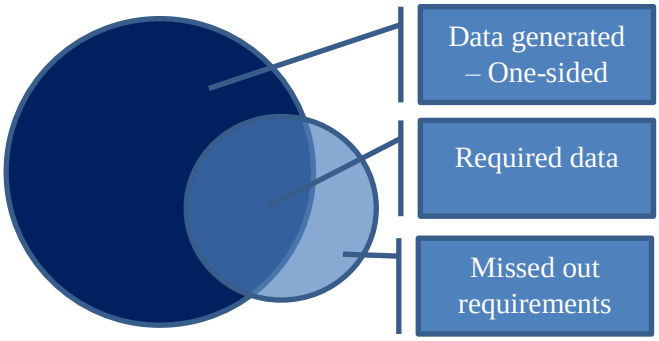


Figure 3. Error of inclusive exclusion

Error 4 – Error of exclusive inclusions

As the term suggests, the error of exclusive inclusions is related to exclusivity (Figure 4). Use of an exclusively specific

term or a Boolean is the principle component of this error. Some other minor contributors to this error can be the use of exclusively specific dialect of language, geography, or exclusively specific exclusion criteria. An error of exclusive inclusion occurs because of extreme caution or critic, and is commonly associated with highly specific key terms used with inadequate Booleans. The search may also exclude synonyms for key terms. In this case, logically, the search phrases are correct and unbiased. Hence, the outcome will have unbiased data that will include positive and negative aspects for all types of possible therapies. However, because of their specificity and exclusion or limited connector's logic, only exact matches are included in the output. Therefore, in some cases, where an antonym of a term consists of the term itself (for example, ST-elevation MI and non-ST-Elevation MI or oriented and non-oriented) may also be excluded. This strategy is useful when the data available are too large, and a very specific search is desired. In other cases, this error has effects similar to that in the error of exclusion.

For example, when searching for non-ST elevation myocardial infarction, a combination of "Coronary Artery Disease" AND ("ECG changes" XOR "ST Elevation") is a specific combination, which will return all coronary artery disease-related literature with electrocardiography (ECG) changes, and will exclude ST-elevation in those ECG changes. However, this search will not return a good literature source because this is most likely to exclude the non-STEMI cases.

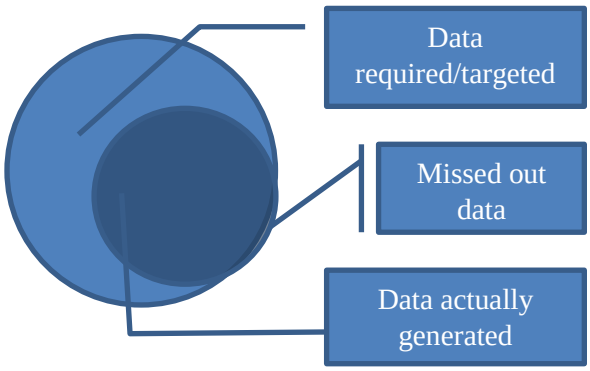


Figure 4. Error of exclusive inclusions

Error 5 – Error of exclusive exclusions (error of limited relevance)

An error of exclusive exclusions is also called the error of limited relevance. This error is a combination of bias and specific exclusiveness. Hence, the search phrases constructed will be biased to only one-sided data trends, and the terms selected will be too exclusive to return sufficient information. Therefore, the output will have a very small relevance to the requirements, and both these sets will actually miss extensive information from the actual available data (Figure 5).

Features of output in this error are linked with relevance. Most of the data generated in this literature search will have limited relevance with the key terms, search phrases, or the purpose of the literature search. Relevant data in the output will be less than one-fifth of the entire output.

Example: "Pacemaker implant" NAND ("Atrial defibrillation" AND "Atrial flutter") taken as a search term for cardiac resynchronization therapy.

Challenges in literature search

On the basis of the type of literature search, two types of challenges exist:

1. Volume-related challenges
2. Relevance-related challenges

Several combinations and permutations of these two challenges can be expressed as practical challenges, namely:

1. Too high data
2. Too low data
3. High volume, low relevance data: Data irrelevant, non-specific, and high on noise
4. Low volume, low relevance data: Non-conclusive search or data; data are specific, but not sufficient to help inferences
5. High value, high relevance data but repetitive outcomes: Monotonous outcomes – no value addition
6. High value, high relevance data, but missing trends and threads: Circular references, cascade of references, or hydra-headed or contradictory results

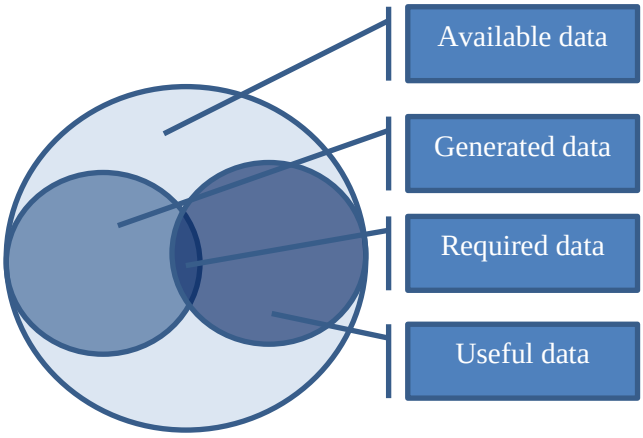


Figure 5. Error of exclusive exclusions

These challenges are discussed in details below.

Challenge 1 – Too high data

This is the most commonly observed challenge in a literature search, especially when the scientist is a novice. The common problem underlying this type of output is an error of inclusion caused by vague, common, general and use of a general database such as Academia, Researchgate, or Google Scholar. In short, this challenge relates to a broad search. For example, the use of the phrase "Event-free Survival" AND "Norfloxacin" returns 4,120 results on Google Scholar; most of the data from the literature search is related to antibiotics in general or similar drugs other than Norfloxacin.

This challenge will increase the quantity of data, and will result in extensive noise and duplication. High volume data makes processing, tabulation, and presentation difficult. In most cases, three-fourth data of this volume is excluded during appraisal. Depending upon further filers of relevance and value addition, about 20% of these data are used in the final documentation for citation and reference purposes.

Challenge 2 – Too low data

This is the second common challenge in the literature search process, when the scientist is too specific and/or has mid-level expertise. The problems with these data extend to two extremes, either the entire data are appraised at once, and the literature search is completed within a short time and with little effort, or the entire process needs to be repeated owing to difficulty in appraisal of the available data.

The common problem with this challenge is the error of exclusion or the use of incorrect logic or an incorrect combination of terms, Boolean, database, or search engine, etc. Depending on further filters of relevance and value addition, the use of data in the final documentation, citation, and reference purposes is unpredictable at the start. In most cases, even if all the data points are appraised from this type of search, their reliability and unbiased nature are uncertain.

Challenge 3 – Irrelevant, non-specific data, and noise

Mostly, data noise is a challenge caused by apprehension of experienced people, who work on a new subject, therapeutic area, or treatment. Especially if a key term has multiple meanings or spelling variation or use in more than one disciplines, its use usually leads to this problem. Replacing specific key terms with non-specific key terms often returns irrelevant or non-specific data. For example, NexGen™ is the name of a stent as well as a knee replacement device. The use of search phrases such as "Nexgen clinical data" may be a medium strength search phrase prima facie; however, when we actually search this in a general database such as Google Scholar, 2300 results are obtained, which include results on knee, stent, and hemophilia studies, constituting a large amount of irrelevant data for a stent.

Challenge 4 – Non-conclusive search of data; data is specific but not sufficient for inferences

When very specific terms are used, we mostly receive information related to the term, but the amount or nature of information is too small to support claims or draw inferences. For example, if we search "Widal test titre 1:80" AND "Antibiotic therapy in low titre enteric pyrexia", no results are returned on Google Scholar, Helios, PubMed, Medscape, and Medline. However, if any word of the phrase or the phrases above are altered, the search returns a large amount of vague data. When the query is split into three parts "Low titre" AND "Antibiotic therapy" AND "Widal test", only 2 articles are returned in the search. Here, the data are too specific but not sufficient to draw inferences or support claims. Rephrasing the search query as "Low titre" AND ("Widal test" AND "Antibiotic") returns 20 articles with specific information on antibiotic therapy, low widal titre, and enteric fever. However, the obtained data cannot be used to obtain conclusions beyond epidemiology and clinical guidelines. When the query is reframed, the search results include extensive noise with no specific information.

Challenge 5 – Monotonous outcomes – no value addition

This challenge is often faced when searching for well-established facts and highly tested treatments or common diseases. For example, the phrase "Aspirin" AND ("Angioplasty" AND "Prophylaxis") return about 8,200 articles, most of which establish the fact that one-year dual antiplatelet therapy and the indefinite use of aspirin are required to prevent stent thrombosis in post-angioplasty care. Even if some well-established and widely accepted trials are obtained from this search, most data will have the same information, usually repetition of the same studies in different contexts such as different timeframes, subsets, and resource. This type of data lacks value addition. Citation of these data is taken only once for its best presentation and source.

Challenge 6 – Circular references, cascade of references, hydra-headed, or contradictory results

Circular references are usually found when two or more contemporary studies are alternately published more than once in different contexts. These studies have two outcome modalities. If all such studies have similar outcomes, the same group is used as a reference to the current claim or statement. However, in certain cases, data from such studies cannot be substantiated, and an expert literature search scientist will refrain from using circular reference studies. The second modality is that these studies have different or contradicting outcomes, the hydra-headed presentation. In such cases, such circular references are better avoided or only the current and best-supported ones should be chosen.

Hydra-headed outcomes are also obtained when terms are used incorrectly and when large research is performed on a condition with the same endpoint and probably different outcomes. Those with minimum experience in literature search on interventional cardiology can easily connect to this situation with the dilemma of choice between coronary artery bypass graft and coronary angioplasty on which various studies are published, and each has different or rather contradictory outcomes. A similar situation was also seen with sympathetic denervation and neuromodulation comparison. A matrix for assessment of causality is presented in Table 1.

Possible corrective and preventive actions and solutions for better literature search.

Preparation of an unbiased literature search protocol

A protocol is among the mandatory requirements of literature search for the clinical evaluation of medical device regulatory activities.[iii] In addition to the regulatory requirements, a literature search protocol should be considered for other types of literature search exercises to maintain a clear and unbiased methodology. A literature search when conducted with the correct protocol is expected to produce the same outputs every time the search is repeated using the same phrases in the same database. Appraisal criteria are other important aspects of the literature search protocol, which need to be correctly set to justify the inclusion, exclusion, and use or rejection of the data points in outputs.

Quality check for search phrases

Quality check at the level of key term, phrase construction, and search engine selection is a critical but often excluded part of the search process. The quality check at these three levels can reduce errors at the root and can save effort. The motive behind a quality check of terms is to avoid bias and ensure appropriate Booleans usage, correct database selection, and optimal use of alternatives.

Automation of literature search

The IT companies with leadership in life-sciences technologies have developed automation of tools that drastically reduce time for literature search and summarization. The use of Summarization, and reporting can be best handled by automation. Algorithm of automation decides its utility and dependability. A good literature search tool provides options for various combinations of key terms. Such a tool provides options of combined access of various databases and provides information on the quality and quantity of output with each key term, search phrase, and database, with a list of unique records. Allows reusability of algorithms and allows ease of customization in different contexts. This reduces the time taken to reach the end-result by avoiding multiple searches. An automated search could provide researchers the desired results based on their pre-specified set criteria, and should significantly decrease the time taken for manual search and filtering appropriate content. This automation will be useful in time-consuming literature search that involves the major task of selecting and extracting desired data. This will reduce the possible bias of inclusive exclusions or exclusive inclusions. Another aspect of automation will be to translate articles from non-English to English and include them in the final selection. This automation of the literature search will eventually reduce the time spent on manual search and would be beneficial in cases in which time is critical in decision making.

Healthcare providers, regulatory agencies, researchers, etc., at times, depend on the latest information in the field to take valuable decisions that could impact public health and/or have large-scale financial implications. Early access to information can ease the decision-making process. Cost saving would be another aspect in automated literature search. Early availability of key information would reduce costs, and less time would be needed for the same task that could lead to better management of resources, including better manpower management. Thus, automation will eventually reduce cost, time, and resource in the whole process. The combination of a good literature search tool and a trained literature search professional can be the best solution to avoid errors and limitation of literature search. Overall literature search strategy is presented in Figure 6.

Acknowledgements

Authors would like to acknowledge Elenka Rodrigues, TATA Consultancy Services, Mumbai, India for editorial assistance.

artificial intelligence and natural language processing based tools with cognitive capabilities provide a near-perfect solution for literature search. For performing literature search for aCER, the tedious job of literature search, appraisal, tabulation,

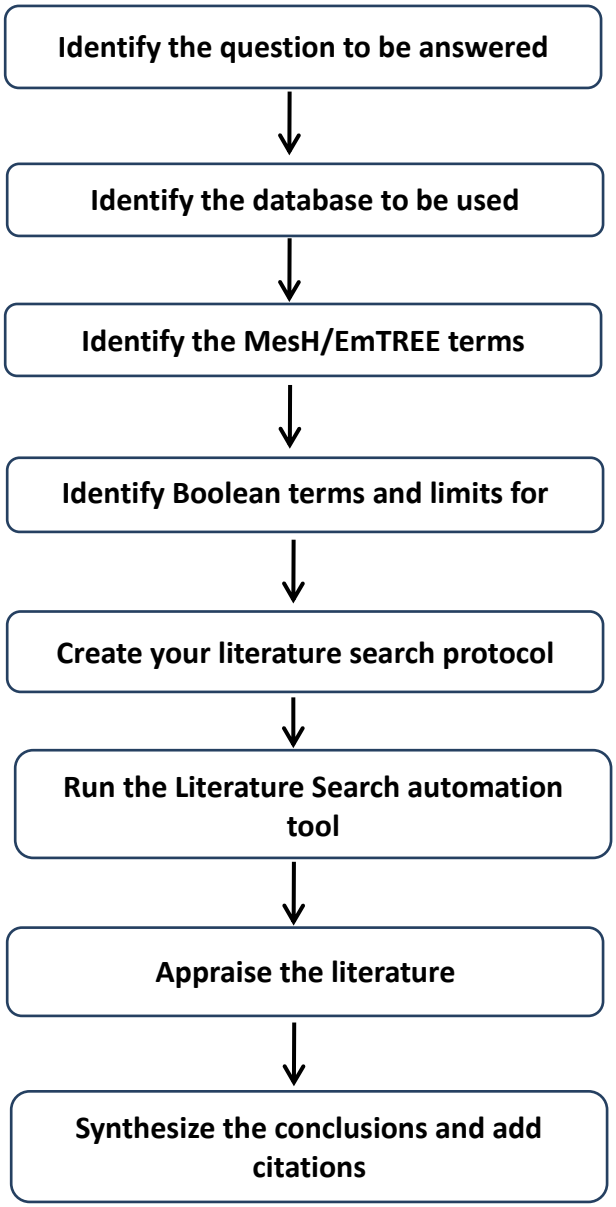


Figure 6. Process flow of literature search

Table 1: Matrix for assessment of causality

Data quantity	Data specificity	Relevance	Inference value	Problem (source of issue)	Possible error	Strategy
High	Specific	Relevant	Conclusive	No problem in most cases Repeating literature or concept	No error Error of inclusion	Handled in appraisal; remove duplicates; refine, if required
			Hydra-headed or self-contradictory	Poor specification of the search criteria Past research has different methods	Error of inclusion	Redefine and recheck the logic
		Irrelevant	Conclusive	Key terms specific, but irrelevant to objective Search logic not correctly set	Error of exclusive inclusion	Rephrase/redefine
			Hydra-headed or contradictory	Use of common synonyms and relevant terms when specific terms are available and use of disease-specific terms Common terms used with a general search engine	Error of inclusion	Use exclusion; refine terms and revise Boolean logic
	Non-Specific	Relevant	Conclusive	Key terms specific but broad, search engine, Obtained results have low scientific value	Error of inclusive exclusion	Redefine, revise logic, rephrase, and change search engine
			Hydra-headed or contradictory	Filters or refining error, incorrect dataset used, too many terms used for the search	Error of inclusion	Redefine, revise logic, rephrase, and refine
		Irrelevant	Conclusive	Vague terms, incorrect terms and phrases Incorrect database choice, and obtained results are usually biased	Error of inclusion	Refine and redefine
			Hydra-headed or contradictory	Vague terms, incorrect terms and phrases Incorrect database choice	Error of inclusion	Discard and re-search
Low	Specific	Relevant	Conclusive	Usually no problem Focused data Biased data must be ruled out	Error of exclusion	Reassign the terms and check for bias
			Hydra-headed or contradictory	Terms and database correct and logic or phrases are incorrect	Error of exclusive inclusion	Redefine
		irrelevant	Conclusive	Incorrect synonyms or logic used Incorrect Boolean used	Error of exclusion	Discard and re-search
			Hydra-headed or contradictory	Too specific terms or phrases and incorrect database	Error of inclusive exclusion	Discard and re-search
	Non-Specific	Relevant	Conclusive	Vague terms or phrases, correct database Wrong Boolean used Use of general terms	Error of exclusive inclusion	Change database

			Hydra-headed or contradictory	Vague terms or phrases, correct database, incorrect Boolean used Use of disease-specific terms instead of treatment-specific terms	Error of exclusive inclusion	Redefine, revise logic, and rephrase
		Irrelevant	Conclusive	Too many filters and multiple logics Use of an incorrect logic for a database	Error of exclusion	Discard and re-search
			Hydra-headed or contradictory	Use of incorrect Boolean logic with too specific terms Use of incorrect logic on database	Error of exclusion	Discard and re-search

AUTHORS

First Author – Ashish Indani, TATA Consultancy Services, Mumbai, India
Second Author – Srinivas Reddy Boreddy, TATA Consultancy Services, Mumbai, India
Third Author – Nadeem Ansari, TATA Consultancy Services, Mumbai, India

Corresponding Author: Dr. Ashish Indani, Tata Consultancy Services, 2nd floor, Tiffany CHS, Hiranandani Estate, Thane (W), India, E-mail: Ashish.indani@tcs.com

i. Garfield, E. and Sher, I. H. (1963), New factors in the evaluation of scientific literature through citation indexing. Amer. Doc., 14: 195–201. doi:10.1002/asi.5090140304
ii. By Mark N.K. Saunders Research Methods For Business Students, 5/e
iii. MedDEV guidelines 2.7.1 rev 3