



## Medical Device Clinical Evaluation Tip 1 - Use of NOT Boolean operator in search strategies

Some people asked me twice these few days about the use of the NOT Boolean operator in search strategies, which prompted me to write this tip.

The use of the NOT operator is not, per se, a source of bias, but it can be if not treated carefully. Why?

There's some reasons, and I will try to outline some here.

The use of the NOT operator can have negative impact (sometimes not even noticed during search strategy design) on the information retrieval. Before explaining this, here is a quick review on information retrieval and clinical evaluation (which is a specific type of systematic review) - more on that in the on How to perform a clinical evaluation of medical devices – Part 1 – Overview and sample of activities - <http://www.medicaldevice.expert/europe/european-commission/medical-device-regulation/how-to-perform-a-clinical-evaluation-of-medical-devices-part-1-overview-and-sample-of-activities/>).

A clinical evaluation will answer one research question that you create based on your device and what you need to know about safety and performance / efficacy / benefits (related to specific Essential / General Safety and Performance Requirements) to show compliance with said Essential / General Safety and Performance Requirements.

The research question needs to be framed/translated using a framework such as PICO to reduce some types of bias and facilitate information retrieval.



### What makes a clinical question well built?

First, the question should be directly relevant to the problem at hand. Next, the question should be phrased to facilitate searching for a precise answer. To achieve these aims, the question must be focused and well articulated for all 4 parts of its 'anatomy':

- The patient or problem being addressed;
- The intervention or exposure being considered;
- The comparison intervention or exposure, when relevant;
- The clinical outcomes of interest.

*Richardson et al, ACP Journal Club, 1995*

The Well-built Clinical Question: A Key to Evidence-based Decisions

These frameworks usually results in concepts that are the basis of the research question (see table 1).

Table 1 - Translation of research question into PICO framework (adapted from Assembling the Pieces of a Systematic Review: A Guide for Librarians and Asking Structured, Answerable Clinical Questions Using the Population, Intervention/Comparator, Outcome (PICO) Framework)

Objective of the research / research question	PICO framework	Research question translation into the PICO framework	Search concepts
Researcher wants to compare the effectiveness of weight loss versus continuous upper airway pressure on blood pressure in sleep apnea patients	P: The patient or problem being addressed I: The intervention or exposure being considered C: The comparison intervention or exposure O: The clinical outcomes	P = Sleep apnea  I = Continuous upper airway pressure  C = Weight Loss  O = Hypertension	Sleep apnea  Continuous upper airway pressure  Weight Loss  Hypertension

However, it's not enough to create a search strategy based on the concepts only, because of the ways datasets and related databases are constructed. For example, If you search only for "sleep apnea", you won't get results for "sleep apnoea", which is another term with the same meaning. In fact, there is always several related terms.

Because of that, you need to perform term harvesting (or pearl harvesting), both from a conceptual (localization) and objective (extraction) standpoints, and both for controlled vocabulary and natural vocabulary.



### Some questions to be asked during term harvesting

What are synonyms for independent variable and dependent variable?

#### **Independent variable:**

- Explanatory variable
- Intervention/comparison, or exposure/comparison
- Exposure variable, or simply exposure
- Predictor, or risk factor

#### **Dependent variable:**

- Response variable
- Outcome variable, or simply outcome

*Speckman et al, PM & R, 2019*

Asking Structured, Answerable Clinical Questions Using the Population, Intervention/Comparator, Outcome (PICO) Framework

One tool to use during term harvesting is to use a term harvest form.

Table 2 - Example term harvesting form for PUBMED (for other databases, there's a need for a specific term harvesting form because of the way each database is indexed and how it works)

Search Fields	PICO Concept 1	PICO Concept 2, 3, ...
<b>Concept</b>	Sleep apnea	
<b>Controlled Vocabulary</b>		
MeSH	Apneas, Central Sleep / Central Sleep Apneas / Sleep Apneas, Central / Apnea, Central / Apneas, Central / Apnea Central / Apneas Central / Apnea, Central Sleep Apnea, Sleep, Central / Sleep Apnea, / Lethal Central / Central Sleep Apnea / Central Sleep Apnea Syndrome / Central Sleep Disordered Breathing / Hypoventilation, Central Alveolar / Alveolar Hypoventilation, Central / Alveolar Hypoventilations, Central / Alveolar Central Hypoventilation / Hypoventilations, Alveolar Central / Ondine Syndrome / Sleep-Disordered Breathing, Central / Breathing, Central Sleep-Disordered / Breathings, Central Sleep-Disordered / Central Sleep-Disordered Breathing / Central Sleep-Disordered Breathings / Sleep Disordered Breathing, Central / Sleep-Disordered Breathings, Central / Central Alveolar Hypoventilation Syndrome / Central Sleep Apnea, Secondary / Secondary Central Sleep Apnea / Sleep Apnea, Newborn, Primary / Primary Sleep Apneas of Newborn / Newborn Primary Sleep Apneas / Central Sleep Apnea, Primary / Primary Central Sleep Apnea	
[MeSH]		
[MeSH:NoExp]		
Medical actions:		
[Medical action]		
Subheadings:		
[Subheading]		

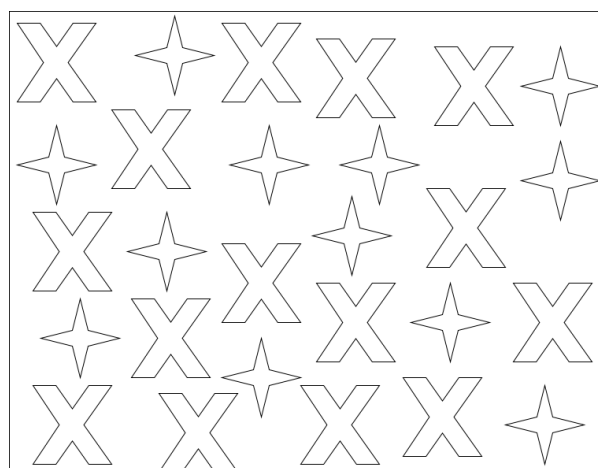


<b>Controlled Vocabulary</b>  Title/Abstract  [Title/Abstract]  OT or TW = Author supplied keywords [Other Term]  [Text Word]		
<b>Natural Vocabulary</b>  Synonyms  UK/US terminology  Medical/laymen's terms  Acronyms/abbreviations  Other narrow search terms  (Consider: phrase searching, proximity operators, truncation, wildcards, truncation, other filters)		
<b>Search result</b>		

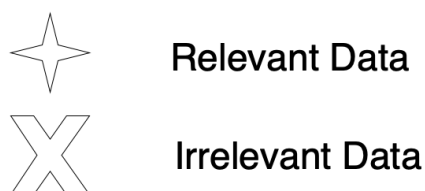
As can be seen from the MeSH example above, PUBMED indexes the concept of “sleep apnea” under several different terms (in fact, the example use the concept of Sleep Apnea, Central”, as Sleep Apnea Syndromes” are divided into central - SLEEP APNEA, CENTRAL, obstructive - SLEEP APNEA, OBSTRUCTIVE, and mixed central-obstructive types).

It's also important to notice that, when you are searching for information (for example, in a database) in a clinical evaluation, you want to get as much as much relevant data as possible (the gold standard is to get ALL relevant data, but for several reasons, this usually do not happen, unless the database is small).

For example, if we think about all the existing data (in a database, in the world, etc.) related to our research question, we can think of the concept of the Universe of Data, with both relevant and irrelevant data related to the research question as can be seen in figure 1.

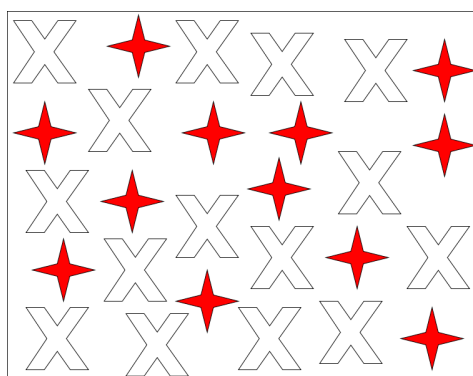


**The Universe of Data**

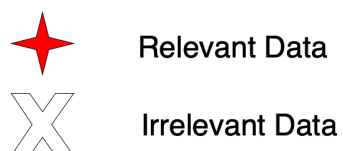


**Figure 1 - The Universe of Data**

What we want, in a clinical evaluation, is to retrieve ALL data, as can be seen in Figure 2.



**The Universe of Data**



**Figure 2 - 100 % retrieval of all relevant data**

However, as mentioned, this is usually not possible.

There are some concepts that are used to evaluate the information retrieval, and two of then most used such concepts are precision and recall (there are several others, such as F-score, mean average precision, etc.)

Precision is usually defined as the number of relevant items from the total number of items a search retrieved.

Numerically, it's defined by this equation:

$$\text{Precision} = \frac{|\{\text{Relevant Data}\} \cap \{\text{Retrieved Data}\}|}{|\{\text{Retrieved Data}\}|}$$

Figure 3 - Equation representing the concept of precision in information retrieval

Visually, it can be seen in Figure 4.

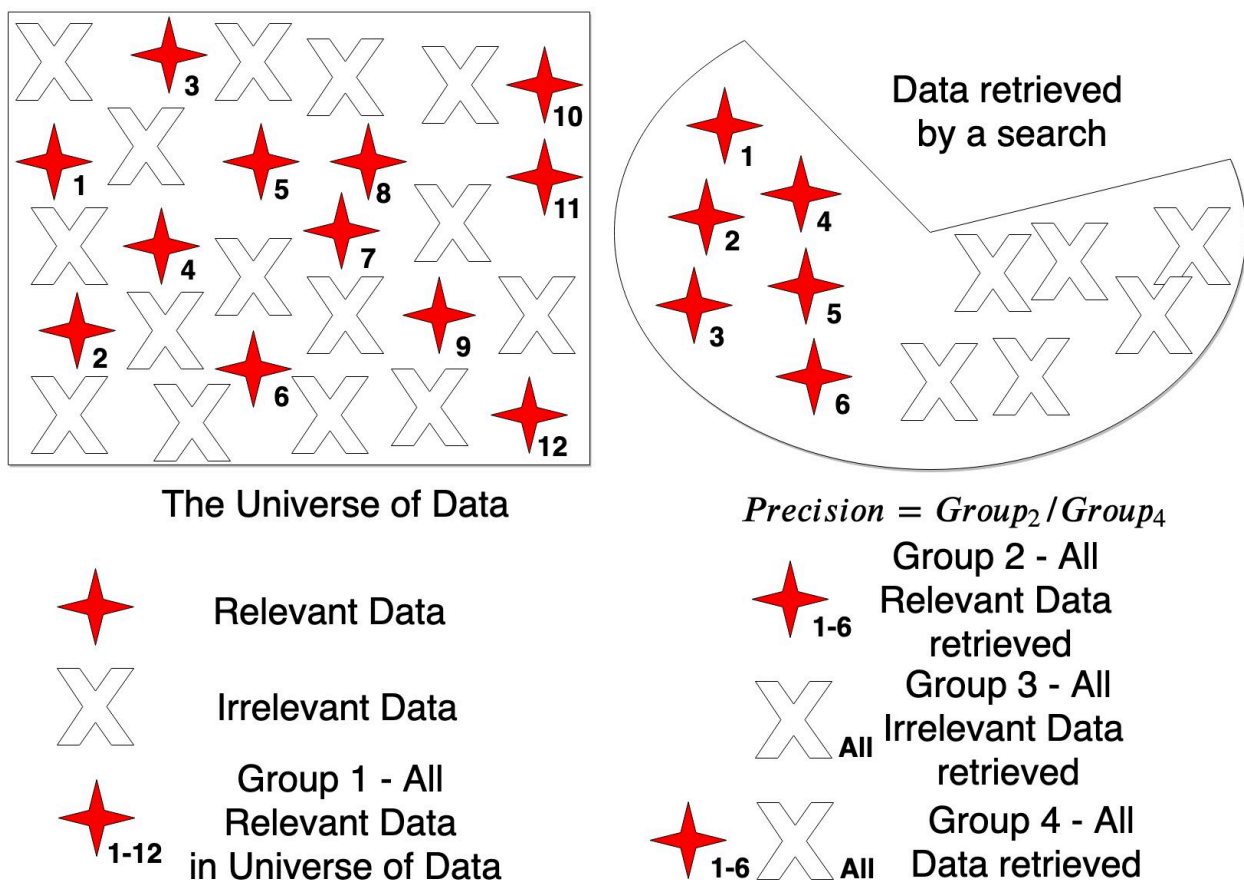


Figure 4 - Visual representation of the concept of precision in information retrieval

Recall is usually defined as total number of relevant items a search retrieved from the total number of relevant items in a database.

Numerically, it's defined by this equation:

$$\text{Recall} = \frac{|\{\text{Relevant Data}\} \cap \{\text{Retrieved Data}\}|}{|\{\text{Relevant Data}\}|}$$

Figure 5 - Equation representing the concept of recall in information retrieval

Visually, it can be seen in Figure 6.

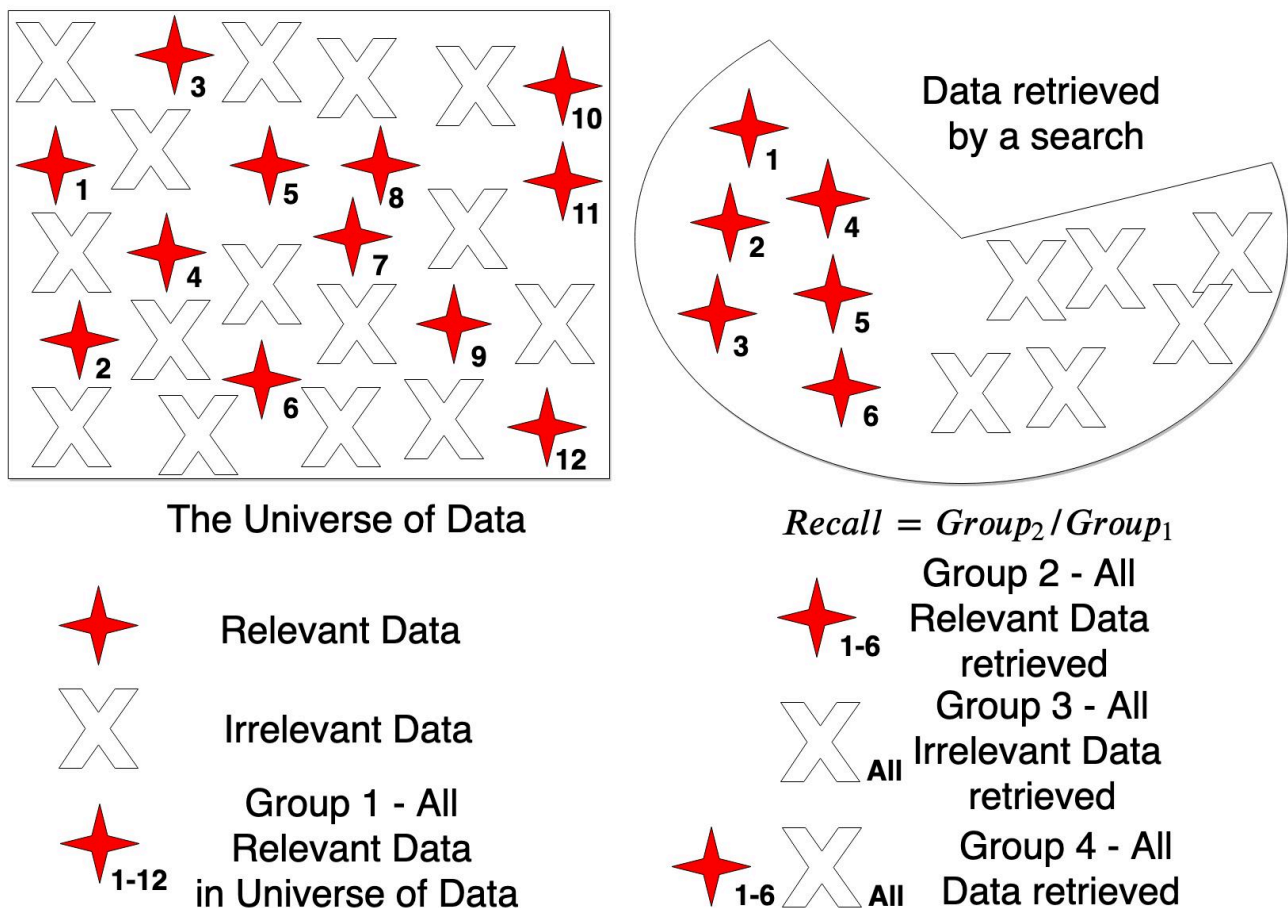


Figure 6 - Visual representation of the concept of recall in information retrieval



Going back to the discussion of the NOT Boolean operator, a quick overview of its use is in Figure 7. The NOT Boolean operator is used to search for terms that appear in the results of Search B (red area in Figure 7) but not in the results of Search A.

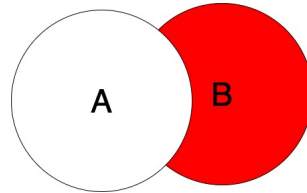


Figure 7 - Results of the use of the NOT Boolean operator (it narrow the search and exclude keywords or subject headings from the search)

The use of the NOT Boolean operator can affect, between other things, the precision and recall of a search strategy. There are some reasons for this.

For example, indexing (both manual and automatic) is not foolproof. It can be the case that the term you removed from the search query using the NOT Boolean operator in fact included relevant data (1 or more), because of a mistake in the indexing process.

Another problem is that when you use the NOT, you are excluding all results with the related terms. If a dataset has both the term and you want and the term you want to exclude, it will exclude it anyway. This can be seen in Figure 8, which also shows why the use of NOT may change precision and recall

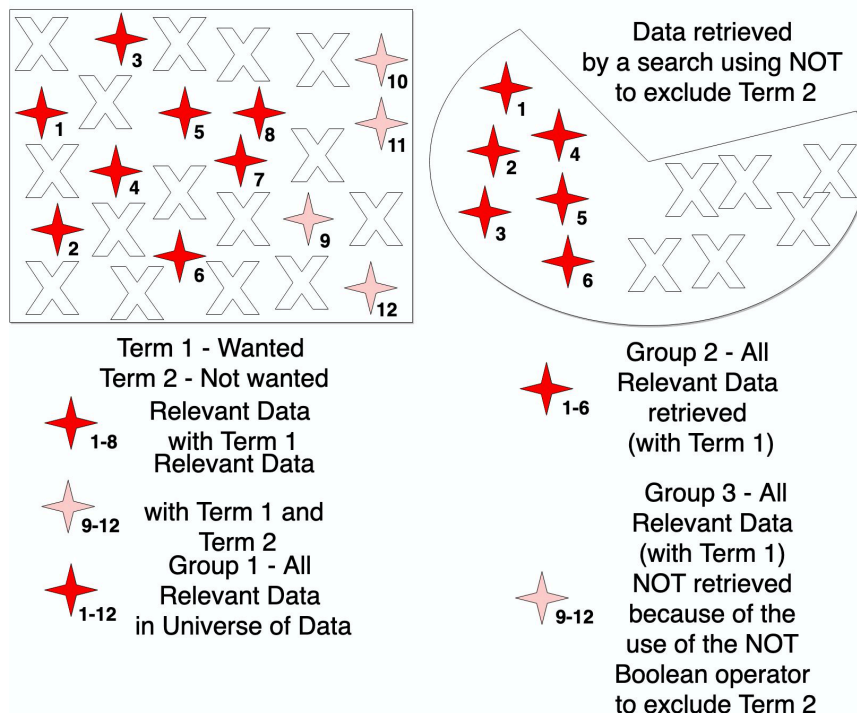


Figure 8 - NOT Boolean operator exclusion of a dataset with both relevant and irrelevant terms



As a final example, another problem that might happen with the use of the NOT Boolean operator particularly in PUBMED (other databases with similar features), is the havoc it can wreck in your search because of automatic term mapping (ATM) or synonym searching, if you do not shut off these features.

Because of these and other possible problems, almost all guidelines suggest the careful use of the NOT Boolean operator.

My personal opinion is even more explicit - I would suggest "NOT" using the NOT Boolean operator if possible on a search strategy. The NOT Boolean operator is great, for example, during the test and calibration of your search strategy, but it would better be off the final strategy.

You can avoid the problems related to non-relevant datasets by screening or, if not caught during screening, by the application of the eligibility criteria. Both cases will require more work, but in my opinion, they are better than missing relevant data (please note that you can still find that you missed relevant data, for example, after performing a manual search for completeness, and in this case you would need to re-calibrate all your search strategies and perform all searches again, but it's better that you take steps to prevent this, like not using the NOT Boolean operator, than to simply wait for this to happen).

IF you choose to use the NOT Boolean operator, here are some suggestions to try and prevent some problems:

- Try to use it only for search strategy testing and calibrating purposes, not for the final search strategy;
- If you use it in your final search strategy, you need to define how much the use of the NOT Boolean operator will affect evaluation parameters such as precision and recall, and make sure that they will not affect them so much in a negative way as to invalidate the results;
- Evaluate how the use of the NOT Boolean operator will create/amplify/change any other bias, and make sure the identified bias is taken care of; and
- Take care with features of databases that might amplify the problems of using the NOT Boolean operator (such as ATM), and preferably turn them off (this is a best practice, anyway, as it gives the search much more control of the results).