

Sensitive versus specific search strategies on PubMed

To the Editor

I read with interest the article on sensitivity of search strategies on PubMed.¹ The authors have presented several PubMed queries to more accurately pickup the articles indexed in the database from Arab countries. Dr. Tadmouri et al, believe that their presented queries are more sensitive.¹

Recalling the definitions of a test (here, our search strategy) sensitivity and specificity,

$$\text{Sensitivity} = \frac{\text{True positive}}{\text{True positive} + \text{false negative}}$$

$$\text{Specificity} = \frac{\text{True negative}}{\text{True negative} + \text{false positive}}$$

we can find that using the NOT Boolean operator, many of the search queries presented¹ (for example those for Palestine (Palestine[affiliation] NOT Jordan[affiliation] NOT TX[affiliation] NOT Lebanon[affiliation] NOT Egypt[affiliation]) or Oman (Oman[affiliation] NOT USA[affiliation] NOT Sweden[affiliation] NOT Stockholm [affiliation]), by excluding the "false-positive" results are indeed more "specific" rather than "sensitive". In fact excluding the "false-positive" results, by any means, will increase the search specificity. On the other hand, when we exclude "false-negative" results (for example queries presented for Bahrain¹ (Bahrain[affiliation] OR Bahrein[affiliation]) or Morocco (Morocco [affiliation] OR Maroc[affiliation]), by using the OR Boolean operator, we will increase the search sensitivity. In more perplexing conditions, however, when we use OR and NOT operators in combination (for example that for Eritrea (Eritrea[affiliation] OR Erythree[affiliation] NOT Italy[affiliation] NOT Milano[affiliation] NOT Germany[affiliation]), the sensitivity and specificity of the query could not be predicted in general. Therefore, the PubMed search queries provided by Tadmouri and Bissar-Tadmouri,¹ though "more accurate," cannot be referred to as "more sensitive" queries, in general.

Farrokh Habibzadeh
NIOC Medical Education and Research Center
PO Box 575
Shiraz 71955
Iran

Reply from the Author

We appreciate the comments of Dr. Habibzadeh on the search strategies we developed, to collect precise data on biomedical research outputs from Arab countries.¹ The reason for which, we used the term "sensitive" to describe our syntaxes is 2 fold. Mathematically, we avoided the use of relatively complex relations such as those suggested by Dr. Habibzadeh, taking into consideration the readers of the journal. Instead, we relied on a much simpler calculation of percent relative error to compare results drawn from non-sensitive and sensitive search strategies. According to the clarification of Dr. Habibzadeh, the use of the Boolean operator NOT, results usually in a withdrawal of false-positive results, thus, a decrease in the number of collected citations. On the other hand, the use of the Boolean operator OR, results in an increase in the number of collected citations. Dr. Habibzadeh referred to the joint use of the Boolean operators NOT and OR in a single search strategy as "perplexing" and its outcome is "unpredictable". However, the results of our suggested strategy indicate a considerable increase in the number of collected citations; hence, a dominance of sensitivity over specificity.¹ This is true for the majority of cases, especially for the most prolific Arab countries regarding biomedical research outputs. On the other hand, the Cochrane Collaboration group, working on the development of search strategies to collect precise results from PubMed,² regularly refers to syntaxes consisting of combinations of all Boolean operators as "highly sensitive strategies."³

Ghazi O. Tadmouri
Nisrine Bissar-Tadmouri
Department of Biology
Bioinformatics Unit
34500 Buyukcekmece
Istanbul
Turkey

References

1. Tadmouri GO, Bissar-Tadmouri N. A major pitfall in the search strategy on PubMed. *Saudi Med J* 2004; 25: 7-10.
2. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *Br Med J* 1994; 309: 1286-1291.
3. Robinson KA, Dickersin K. Development of highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002; 31: 150-153.