

Methodological concerns in Google Trends studies

Observaciones metodológicas en los estudios de tendencias de Google

Carlos Jesús Aragón-Ayala^{1*} <https://orcid.org/0000-0001-9536-2247>

Cender Udai Quispe-Juli² <https://orcid.org/0000-0003-0633-8339>

¹Universidad Nacional de San Agustín de Arequipa. Facultad de Medicina. Arequipa, Perú.

²Universidad Peruana "Cayetano Heredia". Lima, Perú.

*Autor para la correspondencia: caragonay@unsa.edu.pe

Dear editor:

We have read with interest one of the latest published article in your journal.⁽¹⁾ We agree that Google Trends is a useful tool for health research, but some methodological concerns may bias the results of this type of studies. We found in this article serious issues that affect the results and its understandings. We would seek certain clarifications from the authors.

Firstly, the article did not provide the reasoning behind the choice of search input. The selection of search terms should represent what the target population would use for a given concept. Therefore, it is necessary to carry out a pre-validation and careful selection of the search terms,⁽²⁾ either through surveys,⁽³⁾ searching web pages or iteratively testing in Google Trends with the "top related queries".^(4,5) In addition, it is important to show the authors' rationale for search strategy, and especially if it is carried out in several countries.⁽⁶⁾

Secondly, the article did not consider grammatical or language variations of the concepts of interest, despite performing a worldwide analysis. Furthermore, the limitation of the applicability of its results is not mentioned in the discussion or in the conclusion, which would be restricted to English-speaking countries. The search with Google Trends in several geographical areas (or worldwide) should use the "Topic" tool or include translations to the terms into official languages (using the "+" symbol) of the concepts to be searched in "search term" to increase the representativeness of the results.^(2,7,8) For

example, a recent study in Google Trends on the term “wash hands” pre-translated and then evaluated the most searched terms⁽⁹⁾

Thirdly, the authors did not clearly specify whether or not they used quotation marks in their search terms that contained more than one word. This methodological aspect of the search strategy is important for the reproducibility of the results⁽⁸⁾ and it is considered a key performance indicator of methodological quality in Google Trends studies.⁽⁶⁾

Fourthly, the authors do not specify how “Figure 3: “Comparative relative search volume in relation to Internet search trends on COVID-19 symptoms” was made, which shows the comparison of nine search terms related to COVID-19 symptoms.⁽¹⁾ It is well known that Google Trends only allows a comparison of up to a maximum of five, so it is likely that two comparison of groups of terms were downloaded separately and subsequently mixed in the same figure. This should be clarified by the authors because it could affect the authors' conclusions (*the most searched symptom was “fever”, followed by “sore throat” and “cough”*) and the others results.⁽¹⁾ In addition, when trying to reproduce these results (with the same methodological aspects, including “search term” strategy), discrepancies were found with the reported results (Fig.). It is recommended to present these comparisons separately with a same search term for each pair of figures in order to maintain the ranking (in our Figure 1, this *same search term* is “shortness of breath”); or to estimate the absolute value of the searches.⁽⁵⁾

Fifthly, a quantitative parameter was not used for comparison of searches between terms. The authors apparently used a comparative analysis guided by line graphs, which although in this case would not alter their conclusions, is an important aspect to consider especially when there are several search peaks for different terms. A more objective way to make this comparison would be through the area under the curve, the sum of the relative search volumes or their average.

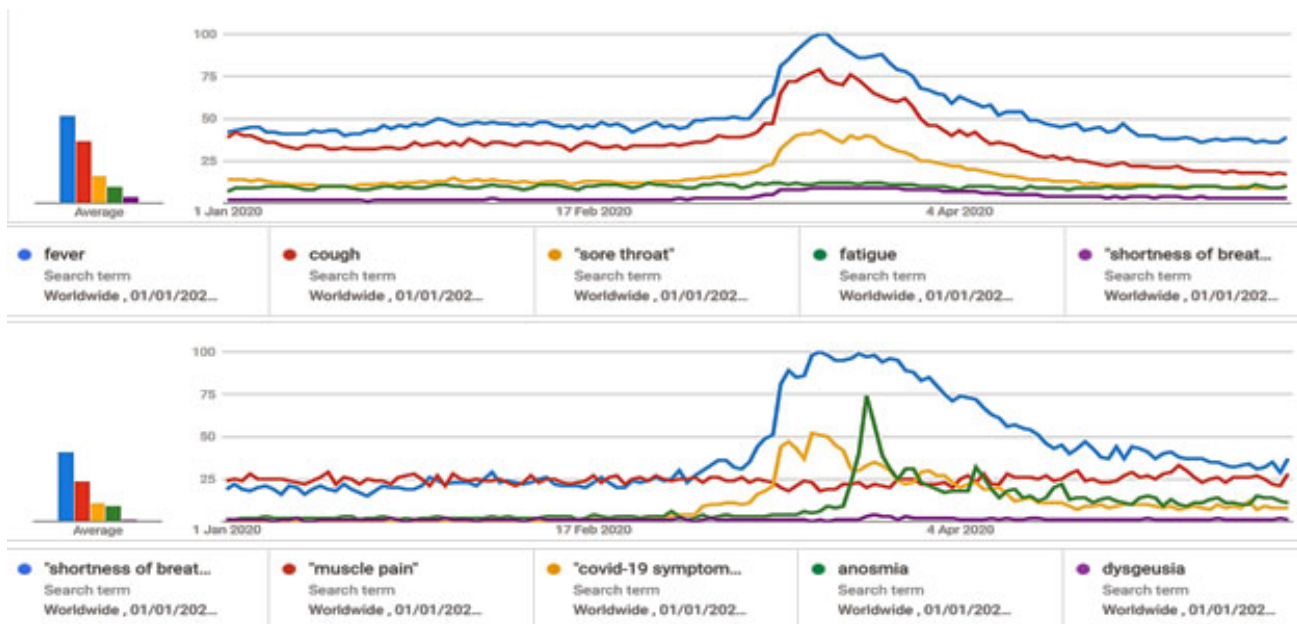


Fig. Covid-19 symptoms sort by average relative search volume. “Shortness of breath” is repeated at the end of the *top figure* and at the beginning of the *bottom figure* to maintain relative comparison.

Finally, one limitation that should be highlighted about the applicability of its results is that some countries, such as China, Iran or Turkey, present disruptions to Google services and therefore a decreased search traffic. There are several causes for these disruptions, such as government-mandated blocks and network outages.⁽¹⁰⁾

To sum up, there are certain issues that should be considered when conducting studies using Google Trends. We hope that the authors will respond to the points presented. This will benefit the scientific community and improve the methodological quality of Latin American research.

Bibliographic references

1. Aquino-Canchari CR, Chávez-Bustamante SG, Caira-Chuquineyra BS. Exploratory analysis of internet search trends during the COVID-19 outbreak. Rev Cubana Inform Cienc Salud [Internet]. 2020 [access: 2020/09/23];31(3). Available from: <http://www.acimed.sld.cu/index.php/acimed/article/view/1631>
2. Mavragani A, Ochoa G. Google Trends in Infodemiology and Infoveillance: Methodology Framework. JMIR Public Health Surveill. 2019 [access: 2020/09/23];29;5(2):e13439. DOI: <http://dx.doi.org/10.2196/13439>

3. Cho S, Sohn CH, Jo MW, Shin SY, Lee JH, Ryoo SM, et al. Correlation between national influenza surveillance data and google trends in South Korea. PLoS One. 2013 [access: 2020/09/23];5;8(12):e81422. DOI: <http://dx.doi.org/10.1371/journal.pone.0081422>
4. Aragón-Ayala CJ, Copa-Uscamayta J, Herrera L, Zela-Coila F, Quispe-Juli CU. Infodemiological study of COVID-19 in Latin America and The Caribbean. MedRxiv; 2020.
5. Zepecki A, Guendelman S, DeNero J, Prata N. Using Application Programming Interfaces to Access Google Data for Health Research: Protocol for a Methodological Framework. JMIR Res Protoc. 2020; 9(7):e16543. DOI: <http://dx.doi.org/10.2196/16543>
6. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. The use of google trends in health care research: a systematic review. PLoS One. 2014;9(10):e109583. DOI: <http://dx.doi.org/10.1371/journal.pone.0109583>
7. Havelka EM, Mallen CD, Shepherd TA. Using Google Trends to assess the impact of global public health days on online health information seeking behaviour in Central and South America. J Glob Health. 2020;10(1):010403. DOI: <http://dx.doi.org/10.7189/jogh.10.010403>
8. Mavragani A, Tsagarakis KP. Predicting referendum results in the Big Data Era. J Big Data. 2019 [access: 2020/09/23];6(1):60. Available from: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-018-0166-z>
9. Lin Y-H, Liu C-H, Chiu Y-C. Google searches for the keywords of “wash hands” predict the speed of national spread of COVID-19 outbreak among 21 countries. Brain Behav Immun. 2020;87:30-2. DOI: <http://dx.doi.org/10.1016/j.bbi.2020.04.020>
10. Google Transparency Report Stories. Tráfico e interrupciones en Google. Google Transparency Report Stories; 2018 [access: 2020/09/23]. Available from: <https://transparencyreport.google.com/traffic/overview?group=REGION>

Conflicts of interest

The authors declare that there are no conflicts of interest.