

Analysis of Google Searches for COVID-19 and its symptoms for predicting disease epidemiology in the United States

Camilla Mattiuzzi¹, Giuseppe Lippi²

¹Service of Clinical Governance, Provincial Agency for Social and Sanitary Services, Trento, Italy

²Section of Clinical Biochemistry, Department of Neuroscience, Biomedicine and Movement, University of Verona, Verona, Italy

To the Editor,

With an ongoing and virtually unstoppable spread of coronavirus disease 2019 (COVID-19), the widespread adoption of epidemiological tools that can be used for anticipating, predicting or monitoring the local viral outbreaks becomes unavoidable (1). In this perspective, the application of infodemiology to managing severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections is emerging as a promising strategy (2,3), provided that appropriate search terms or keywords are used (1), thus preventing the generation of inaccurate or even misleading data on possible epidemiological trends.

With the aim of establishing the most informative clinical correlates of symptomatic SARS-CoV-2 infections, we carried out an electronic search in Google Trends (Google Inc. Mountain View, CA, US), with the search term “COVID-19” together with the most common symptoms reported in patients with SARS-CoV-2 infection, especially in the US, as summarized in Table 1 (4,5). The country option was set to “United States” to increase accuracy of website interrogation, and the search period ranged between the date of the first US case (January 20, 2020) and November 8, 2020. The weekly Google Trends scores for each keyword, which reflect the cumulative Google searches for each specific term during the previous seven days, was downloaded into a Microsoft Excel file (Microsoft, Redmond, WA, United States) for subsequent

statistical evaluation. The correlation between the Google search volume for COVID-19 and the overall number of searches for the other terms during the same period of time was analyzed with both univariate (Pearson’s correlation) and multivariate (multiple linear regression after logarithmic conversion) analyses, where Google Trends score for COVID-19 was entered as dependent variable, whilst the Google Trends scores of terms significantly associated in univariate analysis were entered as independent variables. The statistical analysis was carried out using Analyse-it (Analyse-it Software Ltd, Leeds, UK). The study was conducted in accordance with the Declaration of Helsinki, under the terms of relevant local legislation.

The results of this analysis are shown in Table 1. In univariate analysis a significant association was found between Google search volume for COVID-19 and number of individual Google searches for chills, cough, diarrhea, fever, headache, nausea, sore throat, sputum, smell and taste loss. Altogether the combination of these symptoms explained 88% (95% CI, 76-100%; $p < 0.001$) of total variance of COVID-19 Google searches. No association was observed between Google search volume for COVID-19 and that for dizziness, dyspnea, fatigue, myalgia, rhinorrhea and vomit. In multivariate analysis the number of Google searches for COVID-19 remained significantly associated only with the cumulative number of Google searches for cough and sputum (Table 1).

Table 1. Univariate and multivariate correlations between Google search volume for coronavirus disease 2019 (COVID-19) and cumulative number of searches for the most common symptoms of this disease in the US, between January 20, 2020 and November 8, 2020.

Symptoms	Univariate analysis	Multivariate analysis (β coefficient)
Chills	0.55 (95% CI, 0.29 to 0.74; $p < 0.001$)	1.3; $p = 0.578$
Cough	0.31 (95% CI, 0.00 to 0.56; $p = 0.048$)	7.0; $p < 0.001$
Diarrhea	0.56 (95% CI, 0.30 to 0.74; $p < 0.001$)	5.3; $p = 0.098$
Dizziness	0.14 (95% CI, -0.18 to 0.43; $p = 0.389$)	–
Dyspnea	0.24 (95% CI, -0.08 to 0.51; $p = 0.139$)	–
Fatigue	-0.09 (95% CI, -0.39 to 0.22; $p = 0.558$)	–
Fever	0.57 (95% CI, 0.32 to 0.75; $p < 0.001$)	1.9; $p = 0.508$
Headache	0.77 (95% CI, 0.60 to 0.87; $p < 0.001$)	4.2; $p = 0.366$
Myalgia	0.12 (95% CI, -0.20 to 0.41; $p = 0.464$)	–
Nausea	0.40 (95% CI, 0.10 to 0.63; $p = 0.010$)	3.1; $p = 0.296$
Rhinorrhea	0.16 (95% CI, -0.16 to 0.44; $p = 0.331$)	–
Smell loss	0.56 (95% CI, 0.31 to 0.74; $p < 0.001$)	0.1; $p = 0.948$
Sore throat	0.55 (95% CI, 0.30 to 0.74; $p < 0.001$)	2.2; $p = 0.160$
Sputum	0.55 (95% CI, 0.29 to 0.73; $p < 0.001$)	1.7; $p = 0.017$
Taste loss	0.50 (95% CI, 0.22 to 0.70; $p = 0.001$)	0.1; $p = 0.942$
Vomit	-0.17 (95% CI, -0.45 to 0.15; $p = 0.288$)	–

The results of our analysis of Google search trends show that albeit a number of typical and atypical symptoms are searched by US people who interrogate Google for retrieving information on COVID-19, cough and (to a lesser extent) sputum appear the most correlated with COVID-19 searches. Therefore, monitoring Google search volume for these terms could be a reliable means for predicting the risk of having SARS-CoV-2 infection. Continuous monitoring the number of combined Google searches for chills, cough, diarrhea, fever, headache, nausea, sore throat, smell and taste loss seems instead a valuable strategy for monitoring COVID-19 epidemiology.

Conflicts of interest: Each author declares that he or she has no commercial associations (e.g. consultancies, stock ownership, equity interest, patent/licensing arrangement etc.) that might pose a conflict of interest in connection with the submitted article.

References

- Foley M, O'Neill I, O'Neill B, et al. From bench to bedside - development of an integrated COVID-19 patient flow management system. *J Hosp Infect* 2020;106:211-3.
- Rovetta A, Bhagavathula AS. Global Infodemiology of COVID-19: Analysis of Google Web Searches and Instagram Hashtags. *J Med Internet Res.* 2020 Aug 25;22:e20673.
- Kurian SJ, Bhatti AUR, Alvi MA, et al. Correlations Between COVID-19 Cases and Google Trends Data in the United States: A State-by-State Analysis. *Mayo Clin Proc* 2020;95:2370-2381.
- Argenziano MG, Bruce SL, Slater CL, et al. Characterization and clinical course of 1000 patients with coronavirus disease 2019 in New York: retrospective case series. *BMJ* 2020;369:m1996.
- Yan CH, Faraji F, Prajapati DP, et al. Association of chemosensory dysfunction and COVID-19 in patients presenting with influenza-like symptoms. *Int Forum Allergy Rhinol* 2020;10:806-813.

Correspondence:

Received: 1 December 2020

Accepted: 8 December 2020

Prof. Giuseppe Lippi, Section of Clinical Biochemistry
University Hospital of Verona, Piazzale LA Scuro

37134 Verona, Italy, Tel. +39-045-8124308

Fax. +39-045-8122970

Email: giuseppe.lippi@univr.it