Editorial

Information-seeking about coronavirus in Brazil: a trendy analysis considering online searches

The pandemic due to the novel coronavirus has rapidly spread to all continents within few weeks, and it has been declared a public health emergency by the World Health Organization (WHO) in January 30, 2020.[1] The virus scientifically named SARS-CoV-2 that causes the disease COVID-19 was first reported in Wuhan, Hubei, China in the end of December, 2019.[2]

COVID-19 is a respiratory disease, transmitted by inhalation of droplets disseminated by cough or sneeze, contact with infected oral, nasal and ocular mucosa.[3] Due to the disease exponential nature for transmission, the WHO and Brazilian Ministry of Health (BMH) have proposed the adoption of measures to contain and mitigate the social and health impact of COVID-19 outbreak, such as social distancing, use of personal protective equipment, and non-pharmaceutical strategies, e.g., washing hand, until effective treatment measures become available to prevent transmission and death among individuals.[4]

The wrong measures and the rapid dissemination around the world have caused 9,826 confirmed cases of COVID-19 in January, 85,403 cases in February, 750,890 cases in March, 3,090.445 in April, and in May reaching to 6,120.740 cases. So far, the disease has caused 369,593 deaths (Data collected on May 31, 2020, 4:55pm 15 seconds).[5]

In Brazil, the first COVID-19 case was confirmed in February 26, 2020. The two first infected patients registered were men who lived in city of São Paulo and have return from Italy.[6] Three months after the first case, Brazil become the second country with more cases of the disease in the world, just behind the United States. In terms of number of deaths, Brazil occupies the fourth position, behind the United States, the United Kingdom, and Italy.[5] The analysis of confirmed cases are particularly useful to understand important epidemiological parameters, such as incubation and infections periods, and current outbreaks or probability of occurrence of outbreaks.[6]

In this context, Google Trends appears as data source that can be used to better understand general public interest on the pandemic.[7] Given the serious impact of COVID-19 and the appearance of new methods and approaches to diffuse information and preview behavior in this pandemic, can

be used to explore the relationship between online interest and important events such as the beginning of quarantine and the increase in number of deaths. Currently, infodemiology (epidemiology of information) measures, especially if combined with traditional data, constitutes an integrating part of virus surveillance at regional level.[8]

Considering this, we conducted an objective analysis of online search behavior of those seeking information about the novel coronavirus in Brazil and the disease relationship with incidences of cases, and number of deaths.

Data of internet searches related to coronavirus were obtained from the Google Trends (GT) (https://trends.google.com.br/trends/). GT is a website that presents frequency of searches in Google within specific periods, allowing to observe searching trends in across different geographic regions (i.e., states; countries; or worldwide); topics (health, science, news, travels, among others), and timeframe (specific dates, last 5 years, last week, etc.).[7] GT does not present values in absolute numbers, but its search algorithms normalize results of search volumes from highest searches within specified time period, and the scale ranges from 0 (search volume less than 1% of peak of popularity) to 100 (peak of access) and is presented as relative search volume (RSV).[7]

Google Trends has been used for the analysis of trends about topic related to COVID-19 pandemic in many countries such as Ira,[9] USA,[10,11] Taiwan,[12] France,[13] China,[14] Italy[15]. In addition, comparative approaches have been conducted to determine search behaviors in different continents.[16–18] The methodology employed by GT is described with further details in the website,[7] as well as in other studies.[11,13–15,17–21]

In this study, the profile of explored search considered data from 72 days since the first confirmed case of COVID-19. The analyzed data were from February 26 to May 08, 2020, using the filters of country "Brazil" and topic "health". Data imported as CSV (*Comma-separated values)* files were imported to electronic spreadsheets (MS Excel®).

Search terms were selected based on bibliographical reviews[22], and from empirical and documental observations about predominant expressions used by the BMH and national media to refer to the pandemic. The first terms searched and compared in GT were "coronavirus", "novel coronavirus", and "COVID-19", because the two latter have shown little expressivity in the comparative approach, therefore, we adopted the term "coronavirus". However, it is important to highlight that GT allows a more accurate analysis of search interest by the use of topics of interest, because they consolidated a number of related search terms according to interpretation of Google's algorithm. How terms are written influence their definition, although related terms of topic is already consider in the search box. For this reason, we also performed a study considering the term "coronavírus" with acute accent* that includes also the term "coronavirus" without acute accent, and others related terms.

---

* The acute accent in Brazilian Portuguese designates a stressed vowel. The correct use of acute accent is related to more formal written language.

In the results presented by GT, we observed that the use of term without acute accent presented a higher search volume compared with the no use of acute accent, however, we observed that throughout time, the search term with acute accent gained more volume. Figure 1 shows a comparison between search volumes of "coronavirus" with and without acute accent.
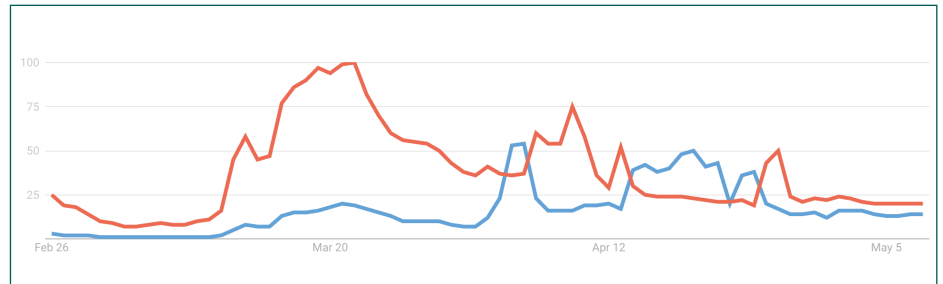

Figure 1. Normalized search volume of the terms coronavirus. Red the term written without acute accent and blue the term written with acute accent

In relation to topic "coronavirus", we observed that the term "coronavirus" (without accent) is the main search term used in this topic. Figure 2 shows search volumes of topic and the term "coronavirus".
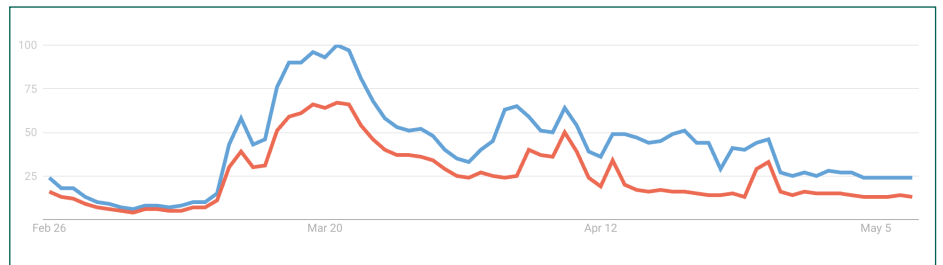

Figure 2. Normalized search volume of term "coronavirus" (red) and for classified searchers by GT belonging to coronavirus (blue) topic

We highlight that due to normalization of data applied by GT algorithm, the maximal search volume in reported graphic will always be 100. We compared the reported volumes by the GT, we observed that peak (100) illustrated in figure 1, corresponded to 67% of all search volume term related to "coronavirus" classified by the Google's algorithm as related to the virus, as is shown in figure 2. That is, in the peak of interest, only around 33% of searches associated by Google's algorithms to "coronavirus" are other, in addition to search the term "coronavirus" without acute accent.

Other evidence about users behavior by using Google is what GT offers are the search of related terms. Figure 3 presents search term related with the topic "coronavirus". Again, result presented with normalized search volumes. Searches conducted using the term with acute accent represented only 44% of searchers compared with "coronavirus" without acute accent. Again, the result is presented with normalized search volumes. The search

using the term with acute accent represented only 44% of searches compared with those using "coronavirus" without accent. We also observed that 21% of searches used the term composed by two separated words: "corona virus". The search term with higher interest was "coronavirus" without acute accent.
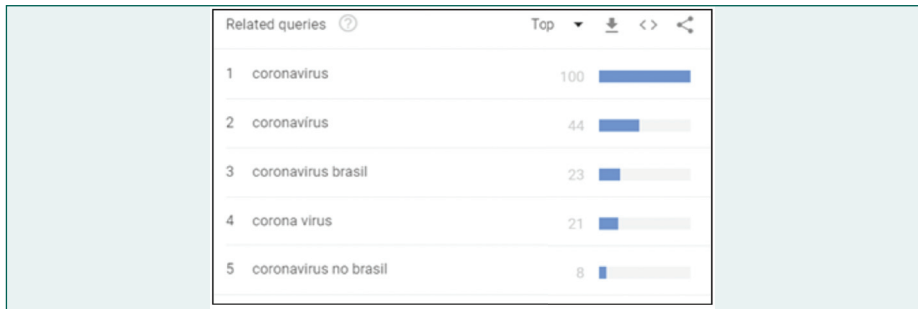

Figure 3. Five most related topics with the searched term "coronavirus"

In relation to related searches, the GT algorithms were also identified by behavioral standards that are classified as emergent interest. Considering the classification by emergent interest, we observed that searches related with coronavirus prevention tips were the one with highest interest, followed by financial aid (government aid) (Figure 4).

Figure 4. Five most significant related search terms with "coronavirus" and identified as of increasing interest by GT

In GT there is the possibility of searching an explanation to phenomena, explore related search term in a specific period. By analyzing the comparison between the term "coronavirus" with or without acute accent, we observed that in April, the term with "accent" is the most used. In the analysis of related terms provide by GT algorithm for the period (Figure 5), we observed that users search conjunctural and technical information about the virus.
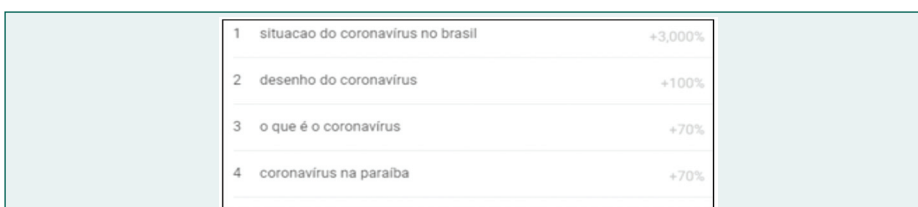

Figure 5. Search related terms to the "coronavirus" that are identified as of increasing interest in GT from April 13, to April 22, 2020 in Brazil

In addition to evidences presented in GT, there is need to consider also in the analysis the occurrence of significant social events that impact the users' behavior. The peak of interest in seeking for information about novel virus coincides with beginning of the quarantine in the main capital of the Brazil.

For this reason, it is plausible to consider that imposition of quarantine and establishment of public emergency situation influenced significantly in the search interest in a number of regions of the country. Users often sought general information about the virus. However, the search term with acute accent became more used from April 2020, which indicate a maturity in information searching on the subject, because the majority of users begin to use in the search the correct writing – with acute accent.

Other interesting phenomena observed related to interest of users on issues is reduction of the search volume of searching through time (Figure 1 and 2). Considering the fact that the search of the term with acute accent has gained importance in relation to the generic term, this might indicate that users become to approach the issue more correctly, searching the term using the most correct wording. However, consideration search of related terms is more plausible to infer that users become educate about the topic by accessed content by searches previously conducted and become to seek information about the virus and disease, passed to follow-up phase with less than ¼ of search volume registered in the peak.

The increasing interest thorough time related to term of seeking for "coronavirus" with acute accent showed a more educated interest by conjunctural updates about the development of the situation, therefore reinforcing the hypothesis of the users interest by the virus evolved in terms of searching for information with the goal to self-education to the stage that users seeking for updates about in relation to the occurrence of the disease.

In addition, in this study we also sought to verify the impact in the interest to search about notification of cases and deaths due to COVID-19. From official data available in BMH, a statistical study was conducted considering the descriptive and inferential analysis of data.

Daily incidences were extracted from official website of the BMH at https://covid.saude.gov.br that currently is updated using data for secretary of health, reflecting advances on cases communication and deaths related to COVID-19 in the country, although some studies emphasize the possibility of data sub-notification.[22]

The IBM SPSS® software, version 24.0, was used for statistical analysis. This software was adopted for graphic production of associations among evolution of notification by the BMH, searches using the term "coronavirus" in GT platform. The Pearson's correlation test was used for inferential analysis of correlation between the number of cases and number of deaths by COVID-19 with search related volume about the disease in the studied period.

Figure 6 includes the graphic with searches volume for the topic "coronavirus" and number of COVID-19 cases, as well as total of registered

deaths in Brazil. Visually, it is possible to perceive that there is no strong correlation between search interest and number of cases or deaths. The correlation between identified cases and deaths is expected by definition, given that most severe cases of the disease end up dying.

We conducted an analysis of correlation using the linear relationship assessment proposed by Pearson. We observed that identified cases and deaths had coefficient correlation of 0.972, however, the correlation of search volume had a negative coefficient of -0.291, which indicated one proportional inversely relationship, i.e., interest decreased although number of cases rose.
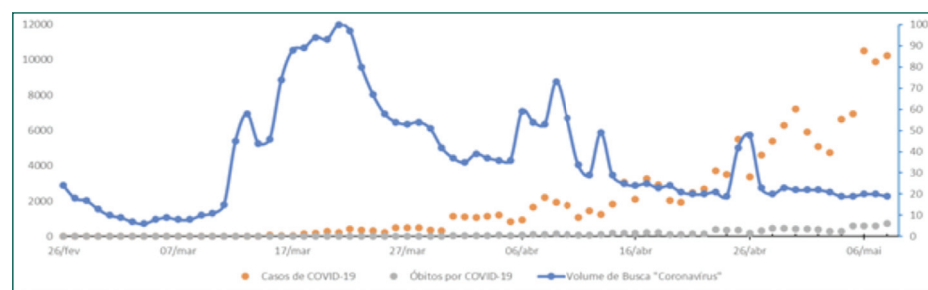

Figure 6. Identified cases of COVID-19 and deaths reported by the BMH, and search volume by the topic "Coronavirus" in GT

It is worthy to highlight that one correlation does not implicate in causality. To identify one correlation is only a signal that some relationship of causality may exist between two variables. The correlation of identified cases with deaths is transitive given that the disease resulted in death to significant amount of identified cases. The relationship of increase of cases with interest checked by searches did not show significant correlation. This verification reinforces that peak of interest in seeking for information about registered situation in searches volume in March was fair, remaining just a little interest for follow-up of the situation, as well as eventual peaks probably motivated by new treatment, although rise in number of cases.

In the relation to ethical aspects, data accessed in GT are of public domain, which disregard the need of submitting the search for an ethical committee. Therefore, this study follows all ethical presupposes of the resolution no. 466/2012 of the Health National Council.

The use of Google Trends was effective to analyze searches of terms related to COVID-19 pandemic. The comparative analysis of the term "coronavirus" with and without acute accent, we observed that the term without "accent" is more frequent, although the remarkable change observed in the users' behavior throughout the time in searchers. We observed that in the beginning of quarantine, users become to search the correct wording of the term in Brazilian Portuguese – with acute accent. This finding may indicate a more educated behavior in terms of literacy, therefore, as well as evidences of this educational improvement and the seeking for information about the situation.

Delmo de Carvalho Alencar (Corresponding author)[1]
*[1]National School of Public Health, Oswaldo Cruz Foundation, Fiocruz,
Rio de Janeiro, RJ, Brazil*
https://orcid.org/0000-0002-6555-7921
E-mail: delmo-carvalho@hotmail.com

Jasilaine Andrade Passos[2]
*[2]National School of Public Health, Oswaldo Cruz Foundation, Fiocruz,
Rio de Janeiro, RJ, Brazil*
https://orcid.org/0000-0001-9595-7822

Ana Raquel Batista de Carvalho[3]
[3]Federal University of Piaui, Teresina, PI, Brazil
https://orcid.org/0000-0001-5287-1084

Aline Raquel de Sousa Ibiapina[4]
[4]Federal University of Piaui, Teresina, PI, Brazil
https://orcid.org/0000-0003-1373-3564

Dárlinton Barbosa Feres Carvalho[5]
[5]Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil
https://orcid.org/0000-0003-3844-0178

Paulo Roberto Vasconcellos-Silva[6]
*[6]National School of Public Health, Oswaldo Cruz Foundation,
Fiocruz,Rio de Janeiro, RJ, Brazil*
https://orcid.org/0000-0003-4646-3580

References

1. World Health Organization (WHO). Statement on the meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). Geneva: WHO; 2020.

2. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet. 2020;395(10223):470–3.

3. Lu CW, Liu XF, Jia ZF. 2019-nCoV transmission through the ocular surface must not be ignored. Lancet. 2020;395(10224):e39.

4. Brasil. Ministério da Saúde. Agência Nacional de Vigilância Sanitária. Orientações Gerais – Máscaras faciais de uso não profissional. Brasília (DF): Ministério da Saúde; 2020.

5. Johns Hopkins Whiting School of Engineering. Center for Systems Science and Engineering. Conornavirus COVID19 Global Cases [Internet]. 2020. [cited 2020 Jun 1]. Available from: https://coronavirus.jhu.edu/map.html

6. Rafael RM, Neto M, Carvalho MM, David HM, Acioli S, Faria MG. Epidemiologia, políticas públicas e pandemia de Covid-19: o que esperar no Brasil? Rev Enferm UERJ. 2020; 28:e49570.

7. Google. Google Trends [Internet]. 2020. [cited 2020 Mai 18]. Available from: https://trends.google.com/trends/

8. Mavragani A. Tracking COVID-19 in Europe: infodemiology Approach. JMIR Public Health Surveill. 2020;6(2):e18941.

9. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, R Niakan Kalhori S. Predicting COVID-19 Incidence through analysis of Google Trends Data in Iran: data mining and deep learning pilot study. JMIR Public Health Surveill. 2020;6(2):e18828.

10. Hong YR, Lawrence J, Williams D Jr, Mainous A 3rd. Population-level interest and telehealth capacity of US hospitals in response to COVID-19: cross-sectional analysis of google search and national hospital survey data. JMIR Public Health Surveill. 2020 Apr;6(2):e18961.

11. Yuan X, Xu J, Hussain S, Wang H, Gao N, Zhang L. Trends and prediction in daily new cases and deaths of COVID-19 in the United States: an internet search-interest based model. Explor Res Hypothesis Med. 2020;5(2):1–6.

12. Husnayain A, Fuad A, Su EC. Applications of Google Search Trends for risk communication in infectious disease management: A case study of the COVID-19 outbreak in Taiwan. Int J Infect Dis. 2020;95:221–3.

13. Kluger N, Scrivener Y. The use of Google trends for acral symptoms during COVID-19 outbreak in France. J Eur Acad Dermatol Venereol. 2020 May 2. doi: 10.1111/jdv.16572.

14. Li C, Chen LJ, Chen X, Zhang M, Pang CP, Chen H. Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. Euro Surveill. 2020;25(10):2000199.

15. Rovetta A, Bhagavathula AS. COVID-19-related web search behaviors and infodemic attitudes in Italy: infodemiological study. JMIR Public Health Surveill. 2020;6(2):e19374.

16. Effenberger M, Kronbichler A, Shin JI, Mayer G, Tilg H, Perco P. Association of the COVID-19 pandemic with Internet Search Volumes: A Google Trends™ Analysis. Int J Infect Dis. 2020 ;95:192–7.

17. Strzelecki A. The second worldwide wave of interest in coronavirus since the COVID-19 outbreaks in South Korea, Italy and Iran: A Google Trends study. Brain Behav Immun. 2020;S0889-1591(20)30551-1.

18. Walker A, Hopkins C, Surda P. The use of google trends to investigate the loss of smell related searches during COVID-19 outbreak. Int Forum Allergy Rhinol. 2020 Apr. https://doi.org/10.1002/alr.22580.

19. Vasconcellos-Silva PR, Carvalho DB, Trajano V, de La Rocque LR, Sawada AC, Juvanhol LL. Using Google Trends Data to study public interest in breast cancer screening in brazil: why not a pink february? JMIR Public Health Surveill. 2017;3(2):e17.

20. Shin SY, Seo DW, An J, Kwak H, Kim SH, Gwack J, et al. High correlation of Middle East respiratory syndrome spread with Google search and Twitter trends in Korea. Sci Rep. 2016;6(1):32920.

21. Poletto C, Boëlle PY, Colizza V. Risk of MERS importation and onward transmission: a systematic review and analysis of cases reported to WHO. BMC Infect Dis. 2016;16(1):448.

22. Platero K, Gomes F. Números estatísticos e realidades: Uma proposta de reflexão sobre a pandemia de Covid-19 no Brasil. Dilemas; 2020. p. 1–11.